DOI:
10.1016/j.jalz.2018.01.017

Document Version
Peer reviewed version

RESEARCH ARTICLE

Elevated DNA methylation across a 48kb region spanning the *HOXA* gene cluster is associated with Alzheimer's disease neuropathology.

Rebecca G. Smith[a], Eilis Hannon[a], Philip L. De Jager[b,c,d], Lori Chibnik[b,c], Simon J. Lott[e], Daniel Condliffe[f], Adam R. Smith[a], Vahram Haroutunian[g,h,i], Claire Troakes[e], Safa Al-Sarraj[e], David A. Bennett[j], John Powell[e], Simon Lovestone[k], Leonard Schalkwyk[l], Jonathan Mill[a,m,*], Katie Lunnon[a,m,*]

[a] University of Exeter Medical School, RILD Building, Royal Devon & Exeter Hospital Campus, Barrack Road, University of Exeter, Devon, UK.

[b] Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Departments of Neurology and Psychiatry, Brigham and Women's Hospital, Boston, MA, USA.

[c] Harvard Medical School, Boston, MA, USA.

[d] Department of Neurology, Columbia University College of Physicians and Surgeons, Columbia University Medical Center, NY, USA

[e] Institute of Psychiatry, Psychology and Neuroscience, King's College London, De Crespigny Park, London, UK.

[f] Queen Mary University of London, London, UK.

[g] Department of Psychiatry, The Icahn School of Medicine at Mount Sinai, New York, USA.

[h] Department of Neuroscience, The Icahn School of Medicine at Mount Sinai, New York, USA.

[i] JJ Peters VA Medical Center, Bronx, New York, USA.

[j] Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA.

[k] Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, UK.

[l] University of Essex, Wivenhoe Park, Colchester, UK.

[m] These authors contributed equally

[*] Corresponding authors:

Katie Lunnon: University of Exeter Medical School, RILD, Barrack Road, University of Exeter, Devon, UK. UK. Tel: + 44 1392 408 298 Email address: k.lunnon@exeter.ac.uk

Abbreviations

Alzheimer's disease (AD); Cerebellum (CER); Differentially methylated position (DMP); Differentially methylated region (DMR); Down syndrome (DS); Entorhinal cortex (EC); Epigenome-wide association study (EWAS); Illumina Infinium Human Methylation 450K BeadChip (450K array); Prefrontal cortex (PFC); Quality control (QC); Reference Sequence (RefSeq); Single nucleotide polymorphism (SNP); Superior temporal gyrus (STG)

Jonathan Mill: University of Exeter Medical School, RILD Building Level 4, Royal Devon and Exeter Hospital, Barrack Rd, Exeter. EX2 5DW. UK. E-mail: j.mill@exeter.ac.uk

Abstract

INTRODUCTION: Alzheimer's disease is a neurodegenerative disorder that is hypothesized to involve epigenetic dysregulation of gene expression in the brain.

METHODS: We performed an epigenome-wide association study to identify differential DNA methylation associated with neuropathology in prefrontal cortex and superior temporal gyrus samples from 147 individuals, replicating our findings in two independent datasets (N=117 and 740).

RESULTS: We identify elevated DNA methylation associated with neuropathology across a 48kb region spanning 208 CpG sites within the *HOXA* gene cluster. A meta-analysis of the top-ranked probe within the *HOXA3* gene (cg22962123) highlighted significant hypermethylation across all three cohorts ($P = 3.11 \text{ X } 10^{-18}$).

DISCUSSION: We present robust evidence for elevated DNA methylation associated with Alzheimer's disease neuropathology spanning the *HOXA* gene cluster on chromosome 7. These data add to the growing evidence highlighting a role for epigenetic variation in Alzheimer's disease, implicating the *HOX* gene family as a target for future investigation.

Keywords

Alzheimer's disease (AD); Braak stage; DNA methylation; Epigenetics; Epigenome-wide association study (EWAS); *HOXA*; Illumina Infinium 450K BeadChip (450K array); Meta-analysis; Neuropathology; Prefrontal cortex (PFC); Superior temporal gyrus (STG)

1. Introduction

Alzheimer's disease (AD), the most common form of dementia, is a progressive neurodegenerative disorder that is making an increasing contribution to the global burden of disease as the population ages [1]. AD pathology is characterized by the accumulation of amyloid (Aβ) plaques and tau tangles, ultimately leading to neuronal cell loss. The neurodegeneration associated with AD is believed to start many decades before clinical onset; during this 'preclinical' phase the plaque and tangle load in the brain increases until a person-specific threshold level is reached and behavioral changes and cognitive impairment become manifest [2-4]. At present there are no disease-modifying treatments available, with existing medications only alleviating certain symptoms of AD. A better understanding of the underlying mechanisms precipitating the onset and progression of pathology is required to enable the design of new, more effective medications.

Increased knowledge about the functional complexity of the genome has led to speculation about the role of epigenetic variation in health and disease, including for neurodegenerative diseases such as AD [5]. Two epigenome-wide association studies (EWAS) of AD [6, 7] recently identified consistent patterns of DNA methylation associated with neuropathology. Of particular interest was replicated evidence for cortex-specific hypermethylation at multiple CpG sites within *ANK1*, although differences at a number of other loci were identified in one, or both studies [8]. One of the previously reported neuropathology-associated differentially methylated positions (DMPs), cg22962123, is located within the *HOXA* gene cluster on chromosome 7 [7]. Here, we present further evidence to support a role for altered DNA methylation in AD-associated neuropathology across an extensive region spanning the *HOXA* cluster.

2. Methods

2.1 Samples and subjects

Our discovery (Mount Sinai) cohort consisted of brain tissue from 147 individuals obtained from the Mount Sinai Alzheimer's Disease and Schizophrenia Brain Bank (http://icahn.mssm.edu/research/labs/neuropathology-and-brain-banking). From the 147 donors, two cortical regions (prefrontal cortex (PFC, N=144) and superior temporal gyrus (STG, N=142)) were used for the purposes of the study. All samples were dissected by trained specialists, snap-frozen and stored at −80°C. Further information about the samples is given in **Supplementary Table 1**. Ethical approval for the project was provided by the University of Exeter Medical School Research Ethics Committee under application number 14/02/041. Genomic DNA was isolated from ~100mg of each dissected brain region using a standard phenol-chloroform extraction protocol and tested for purity and degradation prior to analysis. For replication purposes we used previously published EWAS data collected in two independent cohorts on the Illumina Infinium Human Methylation 450K BeadChip (450K array); i) the "London" (Lunnon *et al*) cohort, consisting of PFC, STG, entorhinal cortex (EC), cerebellum (CER) and pre-mortem blood DNA methylation data from 117 individuals from the MRC London Neurodegenerative Disease Brain Bank [6], and ii) the "ROS/MAP" (De Jager *et al*) cohort, consisting of PFC DNA methylation data from 740 individuals from the Religious Orders Study and the Rush Memory and Aging Project [7]. All samples were assigned a unique code number for the experiment, which was independent of age, gender or diagnosis. This code was used throughout the experiment and analysis.

2.2 Bisulfite treatment and Illumina Infinium BeadArray

500ng genomic DNA was sodium bisulfite converted using the EZ-DNA methylation kit (Zymo Research, Orange, CA, USA) and DNA methylation was subsequently quantified using the 450K array (Illumina, USA) with arrays scanned using an Illumina iScan (software version 3.3.28). Samples were processed by tissue and randomized with respect to age and gender. The Illumina 450K array interrogates >485,000 probes covering 99% of Reference Sequence (RefSeq) genes, with an average of 17 CpG sites per gene region (distributed across promoter, 5'UTR, first exon, gene body and 3'UTR regions). It covers 96% of CpG islands, with additional coverage in island shores and their flanking regions.

2.3 Microarray quality control and data normalization

Initial quality control (QC) of data was conducted using GenomeStudio (version 2011.1) to determine the status of staining, extension, hybridization, target removal, sodium bisulfite conversion, specificity, non-polymorphic and negative controls. Probes previously reported to hybridize to multiple genomic regions or containing a single nucleotide polymorphism (SNP) at the single base extension site were removed from subsequent analyses [9, 10], in addition to the 65 SNPs used for sample identification on the array (total probes removed 72,067). For each probe, DNA methylation levels were indexed by beta values – i.e. the ratio of methylated signal divided by the sum of the methylated and unmethylated signal ($M/M+U$).

2.4 Data analysis

All computations and statistical analyses were performed using R 3.0.2 and Bioconductor 2.13. Signal intensities were imported into R using the *methylumi* package. Initial QC checks were performed using functions in the *methylumi* package to assess concordance between reported and genotyped gender. Non-CpG SNP probes on the array were also used to confirm that both brain regions were sourced from the same individual where expected. Data were pre-processed and quantile normalized using the *dasen* function as part of the *wateRmelon* package (wateRmelon_1.0.3) [11] within the R statistical analysis environment and batch corrected using the *ComBat* package [12]. Array data for each of the tissues was normalized separately and initial analyses were performed separately by tissue. Full Illumina 450K array data was available for the discovery (Mount Sinai) and London (Lunnon *et al*) cohorts and thus we were able to estimate neuronal proportions in the data using the R package *CETS* [13]. For the ROS/MAP (De Jager *et al*) cohort we only had Illumina 450K array data for probes in the *HOXA* region and thus could not calculate neuronal proportions. Therefore, the effects of age, gender and cell type composition were regressed out of the discovery (Mount Sinai) and London (Lunnon *et al*) cohorts, whilst the effects of age and gender only were regressed out of the ROS/MAP (De Jager *et al*) cohort before subsequent analysis. For identification of differentially methylated positions (DMPs) specifically altered with respect to neuropathological measures of AD, we performed a quantitative analysis in which samples were analyzed separately in each brain region using linear regression models with respect to Braak stage, with probes ranked according to *P* value. The genic location of identified DMPs was annotated by GREAT annotation [14]. We have previously established the multiple testing threshold (experiment-wide significance) for EWAS data generated on the Illumina 450K array as $P < 2.2 \times 10^{-7}$ [15]. In brief, in this previous study 5000 permutations were performed repeating a linear regression model for randomly selected groups of cases and controls (N=675). For each permutation, *P* values from the EWAS were saved and the

minimum identified. Across all permutations the fifth percentile was calculated to generate the 5 % alpha significance threshold, which was deemed to be $P < 2.2 \times 10^{-7}$. To identify differentially methylated regions (DMRs), we identified spatially correlated *P* values in our data using the Python module *comb-p* to group ≥3 spatially correlated CpGs in a 500-bp sliding window [16]. The *coMET* package was used to identify regional co-methylation patterns and regional EWAS results [17]. Fisher's combined *P* value analysis was performed in the *MetaDE* package [18] and meta-analysis on correlation and case control status performed with the *meta* package [19] within R [20]. Data is available for the discovery (Mount Sinai) cohort within GEO under accession number GSE80970. The discovery (Mount Sinai) EWAS data-set has been previously used to validate the top 100 DMPs nominated in a previously published EWAS [6]. As such, we have not sought to replicate these top 100 DMPs in the current study.

3. Results

3.1 Hypermethylation associated with AD neuropathology is observed in a region spanning 48kb across the HOXA gene cluster in human cortex

Our primary analyses focused on matched PFC and STG tissue from 147 individuals (**Supplementary Table 1**). We used the 450K array to first quantify DNA methylation in the PFC and identify DMPs associated with Braak score, a standardized measure of neurofibrillary tangle burden determined at autopsy, controlling for age, gender and estimated neuronal cell proportion. We identified ten experiment-wide significant ($P < 2.2 \times 10^{-7}$) DMPs (**Table 1A** and **Figure 1A**), with 78 DMPs associated with Braak stage at a more relaxed threshold of $P < 1 \times 10^{-5}$ (**Supplementary Table 2**). Of these 78 DMPs, nine were located in the *HOXA* gene cluster on chromosome 7, most notably in the vicinity of *HOXA3*,

with one *HOXA* DMP reaching experiment-wide significance (cg22962123: $P = 1.2 \times 10^{-7}$). We next used a sliding window approach (*comb-p [16]*) to identify spatially correlated regions of differential DNA methylation associated with neuropathology; **Table 1B** lists differentially methylated regions (DMRs) spanning at least three probes with a window size of 500bp and a Sidak-corrected *P* value <0.05. We identified six closely-located DMRs within the *HOXA* gene region, with the most significant DMR in the *HOXA* region spanning seven probes in a 364bp region within intron 1 of *HOXA3* (**Figure 1B**; Sidak-corrected $P = 1.19 \times 10^{-9}$). Of note, we observed an extended region of neuropathology-associated hypermethylation spanning 48,754 base pairs (bp) from upstream of the *HOXA2* gene to the *HOXA6* gene and covering 208 Illumina 450K array probes (**Figure 1C**). Given that DNA methylation at nearby CpG sites can be highly correlated [21], we visualized co-methylation patterns between CpG sites within *HOXA3* using *coMET* [17] and observed highly correlated patterns of DNA methylation between CpG sites in this extended region (**Supplementary Figure 1**). We next sought to test whether neuropathology-associated DNA methylation patterns across this 48,754bp region were specific to the PFC, using the Illumina 450K array to profile STG samples from the same individuals. In total, seven probes in the region demonstrated significantly increased DNA methylation after correcting for 208 tests ($P < 2.4 \times 10^{-4}$), with the top PFC DMP (cg22962123) being similarly hypermethylated with respect to Braak stage (**Figure 1D**; PFC: R = 0.36, $P = 1.2 \times 10^{-7}$; STG: R= 0.28, $P=2.78 \times 10^{-4}$). There was an overall consistent pattern of effect sizes across both brain regions for the 208 probes in the *HOXA* neuropathology-associated region (**Figure 1E**; R = 0.76, $P = 2.66 \times 10^{-40}$).

3.2 Cortical neuropathology-associated hypermethylation in HOXA3 is observed in independent study cohorts

We next sought to replicate the observation of neuropathology-associated hypermethylation across these 208 probes in two independent, previously published data-sets. First we examined the "London" (Lunnon *et al* [6]) data-set, comprising Illumina 450K array data generated using matched PFC, STG, EC, CER and pre-mortem blood samples obtained from 117 donors (described in [6]) (**Supplementary Table 1**). We observed a similar pattern of Braak-associated DNA methylation across this 208 probe region in the replication cohort in both the PFC (**Figure 2A**) and STG (**Supplementary Figure 2**), with a highly correlated effect size between cohorts in both brain regions (PFC: **Figure 2B**; R = 0.74, P = 2.27 X 10^{-37}; STG: **Supplementary Figure 3**; R = 0.68, P = 1.87 X 10^{-29}); 15 and six probes reaching our corrected significance threshold (P < 2.4 X 10^{-4}) in the PFC and STG respectively. In contrast, no probes in this region reached the corrected significance threshold in the EC (**Supplementary Figure 4**) although the effect size was still correlated (R = 0.41, P = 1.23 X 10^{-9}). Similarly no probes reached the significance threshold in the CER (**Supplementary Figure 5**), or in pre-mortem whole blood collected in a subset (N=57) of the same individuals (**Supplementary Figure 6)**, with no correlation of effect sizes in either the CER (R = 0.03, P = 0.639) or blood (R = 0.11, P = 0.138). This indicates that the association may be specific to only particular regions of the cortex.

We subsequently assessed this region in the "ROS/MAP" (De Jager *et al*) data-set comprising of 740 PFC samples profiled on the Illumina 450K array (as described in Ref [7]) (**Supplementary Table 1**) observing a similar pattern of effects with highly-significant neuropathology-associated hypermethylation across probes in the *HOXA* genic region (**Figure 2C**), and a significant correlation of effect size with the same 208 probes in the PFC in the discovery cohort (**Figure 2D**; R = 0.80, P = 2.39 X 10^{-48}). A Fisher's combined P value of DNA methylation differences across this region in all three PFC datasets confirmed a clearly defined region of significant neuropathology-associated elevated DNA methylation,

with many individual DMPs passing the threshold for experiment-wide significance (**Figure 2E**), and a consistent pattern of effects across the three cohorts (**Supplementary Figure 7**). The most significant DMP identified within the *HOXA3* gene in our discovery cohort (cg22962123; **Table 1A**) was also the most significant DMP in our Fisher's combined P value analysis ($P = 1 \text{ X } 10^{-20}$). A meta-analysis comparing Braak 0 to Braak VI demonstrated increased DNA methylation with respect to Braak stage across all cohorts in the PFC (**Figure 2F**; $P = 3.11 \text{ X } 10^{-18}$). Together, our data suggests that DNA hypermethylation across the extended *HOXA* gene region is robustly associated with AD-related neuropathology in both the PFC and STG, with the strongest effects in the vicinity of *HOXA3*.

4. Discussion

We identified an extended region of elevated DNA methylation in the *HOXA* gene cluster that is associated with AD neuropathology, with consistent effects seen across three independent post-mortem brain sample cohorts. Although one previous study had demonstrated differential methylation at a single CpG within the *HOXA* gene cluster [7], and another identified a DMR spanning seven CpG sites [6], this represents the first study to illustrate that hypermethylation in this region extends to 208 DMPs, spanning approximately 48.7Kb. Differential DNA methylation in the *HOXA* gene cluster has been previously reported in blood collected from Down syndrome (DS) individuals [22], which is interesting given that many DS individuals develop AD resulting from an additional copy of the *APP* gene due to trisomy on chromosome 21. The DS study demonstrated differential DNA methylation in 20 probes largely located within *HOXA2*. Of note, 17 of these probes were significantly hypermethylated in the PFC in our discovery (Mount Sinai) cohort. However none were differentially methylated in pre-mortem blood in the London (Lunnon *et al*)

cohort. In the context of other neurodegenerative disorders, one study that investigated microRNAs (miRNAs) targeting *HOX* genes in Huntington's disease (HD) demonstrated increased levels of miRNAs related to *HOXA5*, *HOXA10*, *HOXA11*, *HOXA-11AS*, *HOXA13* and *HOTAIRM1* in the prefrontal cortex in HD [23]. Although *HOX* genes encode potent transcription factors that play a critical role in embryonic development [24], a recent study in *Drosophila* also highlighted a potent protective function for *Hox* genes in neurons, implicating a role in neuro-protection [25] Interestingly, this study also highlighted how *Hox* genes act to maintain expression of the ankyrin locus, an important observation given our previous finding of altered DNA methylation in *ANK1* in AD [6]. Indeed, to further explore this hypothesis we examined the correlation between DNA methylation levels at the most significant *HOX* probe identified in the current study (cg22962123) with the two *ANK1* DMPs that we previously identified to be associated with AD neuropathology (cg11823178 and cg05066959) [6, 7] in the PFC, identifying a significant correlation with both *ANK1* probes (cg11823178: R = 0.24, P = 5.15 X 10^{-10}; cg05066959: R = 0.20, P = 2.93 X 10^{-8}). Although this correlation could reflect the association between both *HOXA3* and *ANK1* probes with Braak stage, it could highlight a novel physiological mechanism, particularly as we still observed significant hypermethylation (P = 1.67 X 10^{-5}) at our top *HOXA* probe (cg22962133), when controlling for levels of DNA methylation in the top *ANK1* probe (cg11823178). Looking to the future, analyses of gene expression levels should be performed to facilitate the interpretation of the DNA methylation differences we observe in *HOXA*. To conclude, this study provides further evidence for altered epigenetic processes in the pathophysiology of AD and suggests that further work on the neuroprotective functions of *HOX* genes is warranted.

Acknowledgements

This work was funded by NIH grant R01 AG036039 to JM, Alzheimer's Society grant AS-PG-14-038 to KL, Alzheimer's Association grant NIRG-14-320878 to KL and a grant from BRACE (Bristol Research into Alzheimer's and Care of the Elderly) to KL. Brain banking and neuropathology assessments for the discovery cohort from the Mount Sinai Alzheimer's disease and Schizophrenia Brain Bank was supported by NIH grants AG02219, AG05138 and MH064673 and the Department of Veterans Affairs VISN3 MIRECC. Brain banking and neuropathological assessment for the London (Lunnon *et al*) cohort was provided by The London Neurodegenerative Diseases Brain Bank, which receives funding from the MRC and as part of the Brains for Dementia Research (BDR) programme, jointly funded by Alzheimer's Research UK and Alzheimer's Society. The ROS/MAP (De Jager *et al*) cohort was supported by the National Institutes of Health grants: R01 AG036042, R01AG036836, R01 AG17917, R01 AG15819, R01 AG032990, R01 AG18023, RC2 AG036547, P30 AG10161, P50 AG016574, U01 ES017155, KL2 RR024151, K25 AG041906-01.

References

[1] Prince M, Guerchet M, Prina M. The Global Impact of Dementia 2013–2050. Alzheimer's Disease International (ADI). 2013.

[2] Blennow K, de Leon MJ, Zetterberg H. Alzheimer's disease. Lancet. 2006;368:387-403.

[3] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2011;7:280-92.

[4] Jack CR, Jr., Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. Lancet neurology. 2010;9:119-28.

[5] Lunnon K, Mill J. Epigenetic studies in Alzheimer's disease: current findings, caveats, and considerations for future studies. American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics. 2013;162B:789-99.

[6] Lunnon K, Smith R, Hannon E, De Jager PL, Srivastava G, Volta M, et al. Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. Nature neuroscience. 2014;17:1164-70.

[7] De Jager PL, Srivastava G, Lunnon K, Burgess J, Schalkwyk LC, Yu L, et al. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. Nature neuroscience. 2014;Sep;17:1156-63.

[8] Lord J, Cruchaga C. The epigenetic landscape of Alzheimer's disease. Nature neuroscience. 2014;17:1138-40.

[9] Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics : official journal of the DNA Methylation Society. 2013;8:203-9.

[10] Price ME, Cotton AM, Lam LL, Farre P, Emberly E, Brown CJ, et al. Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. Epigenetics Chromatin. 2013;6:4.

[11] Pidsley R, Wong CCY, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC genomics. 2013;14:293.

[12] Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28:882-3.

[13] Guintivano J, Aryee M, Kaminsky Z. A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. Epigenetics : official journal of the DNA Methylation Society. 2013;8.

[14] McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. Nature biotechnology. 2010;28:495-501.

[15] Hannon E, Dempster E, Viana J, Burrage J, Smith AR, Macdonald R, et al. An integrated genetic-epigenetic analysis of schizophrenia: Evidence for co-localization of genetic associations and differential DNA methylation. Genome biology. 2016.

[16] Pedersen BS, Schwartz DA, Yang IV, Kechris KJ. Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. Bioinformatics. 2012;28:2986-8.

[17] Martin TC, Yet I, Tsai PC, Bell JT. coMET: visualisation of regional epigenome-wide association scan results and DNA co-methylation patterns. BMC bioinformatics. 2015;16:131.

[18] Wang X, Li J, Tseng GC. MetaDE: Microarray meta-analysis for differentially expressed gene detection. 2012.

[19] Schwarzer G. meta: General Package for Meta-Analysis. 2015.

[20] Wang X, Kang DD, Shen K, Song C, Lu S, Chang LC, et al. An R package suite for microarray meta-analysis in quality control, differentially expressed gene analysis and pathway enrichment detection. Bioinformatics. 2012;28:2534-6.

[21] Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome biology. 2011;12:R10.

[22] Bacalini MG, Gentilini D, Boattini A, Giampieri E, Pirazzini C, Giuliani C, et al. Identification of a DNA methylation signature in blood cells from persons with Down Syndrome. Aging. 2015;7:82-96.

[23] Hoss AG, Kartha VK, Dong X, Latourelle JC, Dumitriu A, Hadzi TC, et al. MicroRNAs located in the Hox gene clusters are implicated in huntington's disease pathogenesis. PLoS genetics. 2014;10:e1004188.

[24] Krumlauf R. Hox genes in vertebrate development. Cell. 1994;78:191-201.

[25] Friedrich J, Sorge S, Bujupi F, Eichenlaub MP, Schulz NG, Wittbrodt J, et al. Hox Function Is Required for the Development and Maintenance of the Drosophila Feeding Motor Unit. Cell reports. 2016;14:850-60.

FIGURE 1

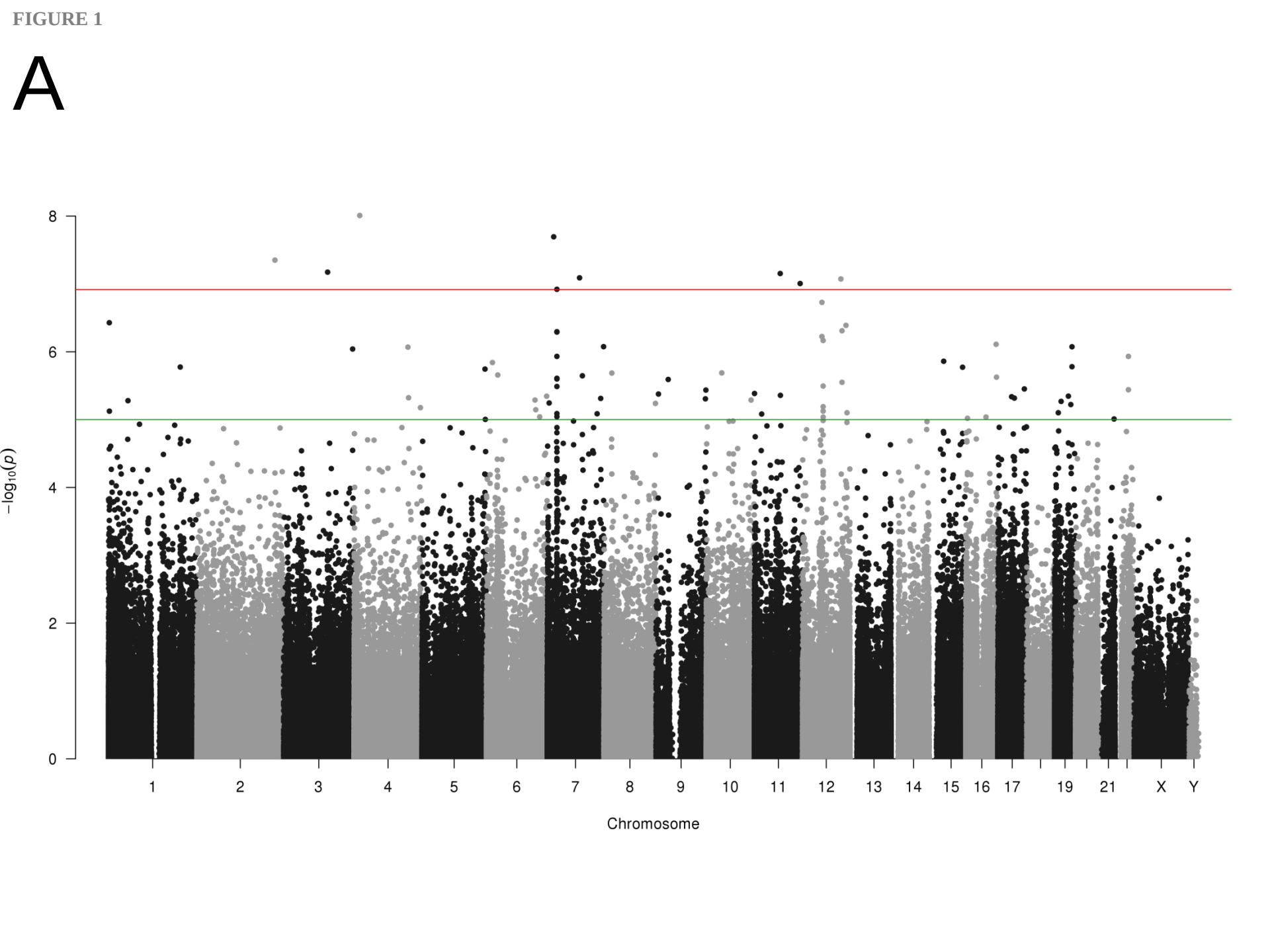

A
8
6
4
2
0
−log10(p)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
19
21
X
Y
Chromosome

FIGURE 1

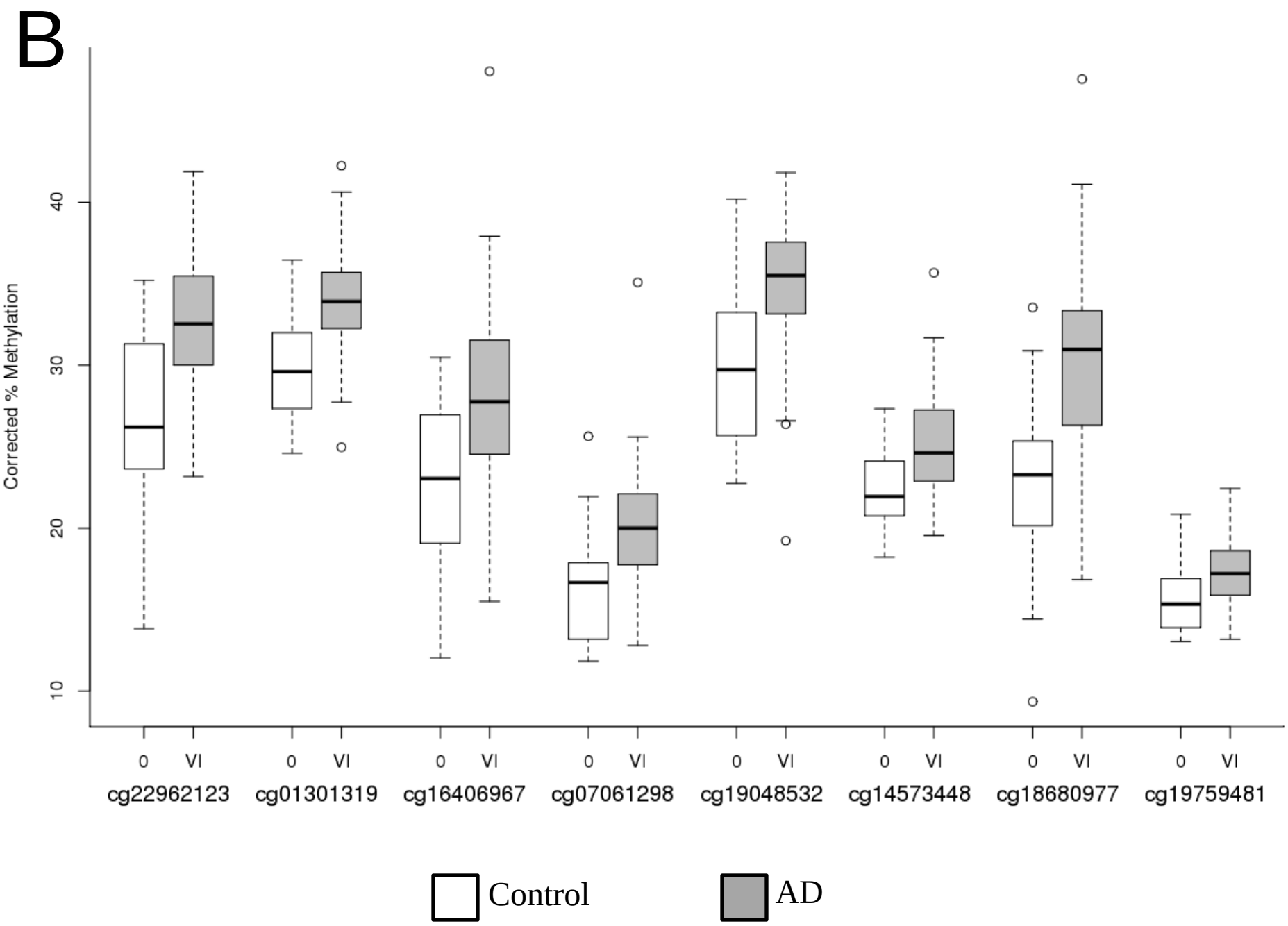

B
Corrected % Methylation
10
20
30
40
0
VI
cg22962123
cg01301319
cg16406967
cg07061298
cg19048532
cg14573448
cg18680977
cg19759481
Control
AD

FIGURE 1

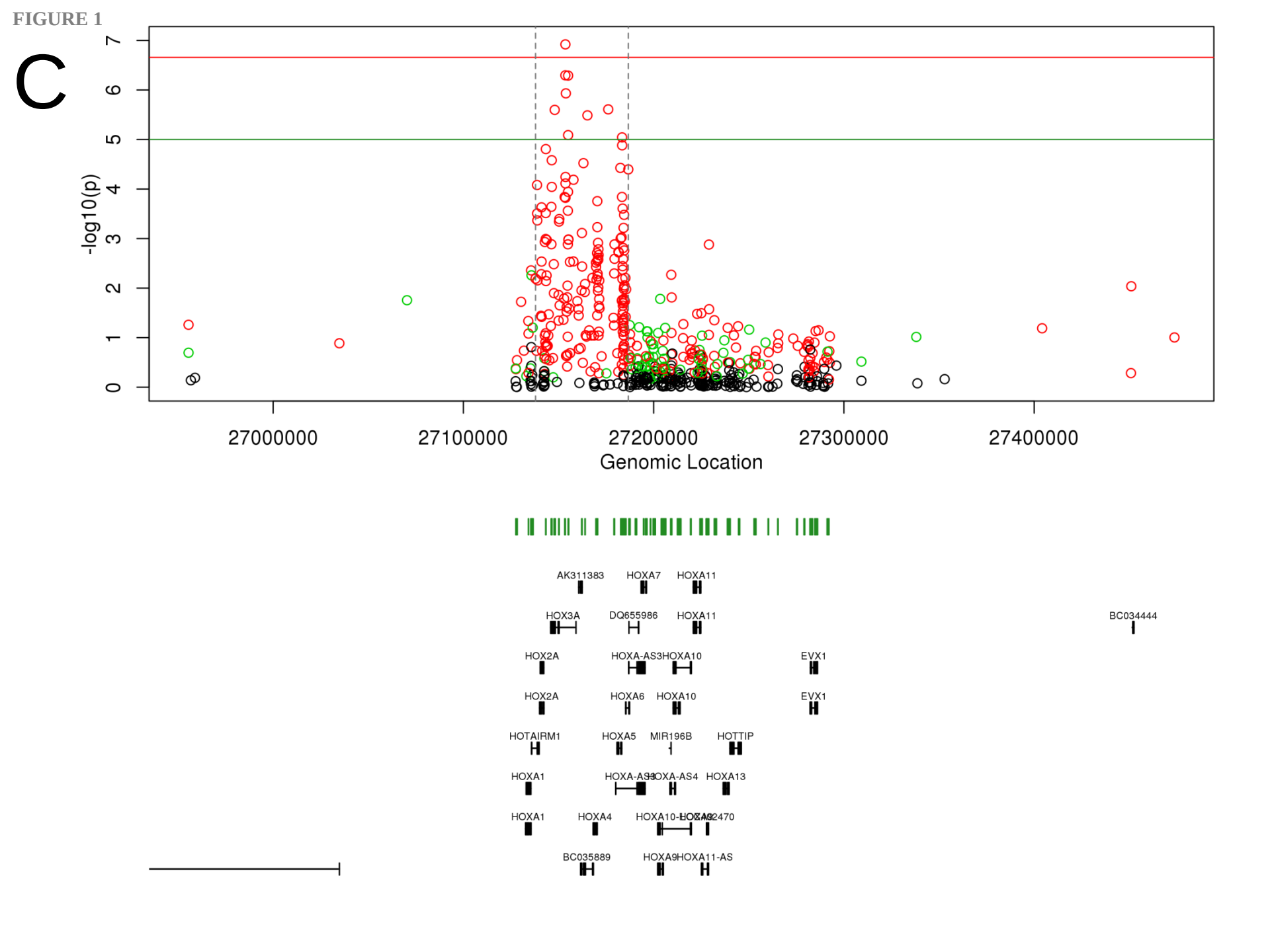

C
-log10(p)
0
1
2
3
4
5
6
7
27000000
27100000
27200000
27300000
27400000
Genomic Location
AK311383
HOXA7
HOXA11
HOX3A
DQ655986
HOXA11
BC034444
HOX2A
HOXA-AS3
HOXA10
EVX1
HOX2A
HOXA6
HOXA10
EVX1
HOTAIRM1
HOXA5
MIR196B
HOTTIP
HOXA1
HOXA-AS3
HOXA-AS4
HOXA13
HOXA1
HOXA4
BC035889
HOXA9
HOXA11-AS

FIGURE 1

D

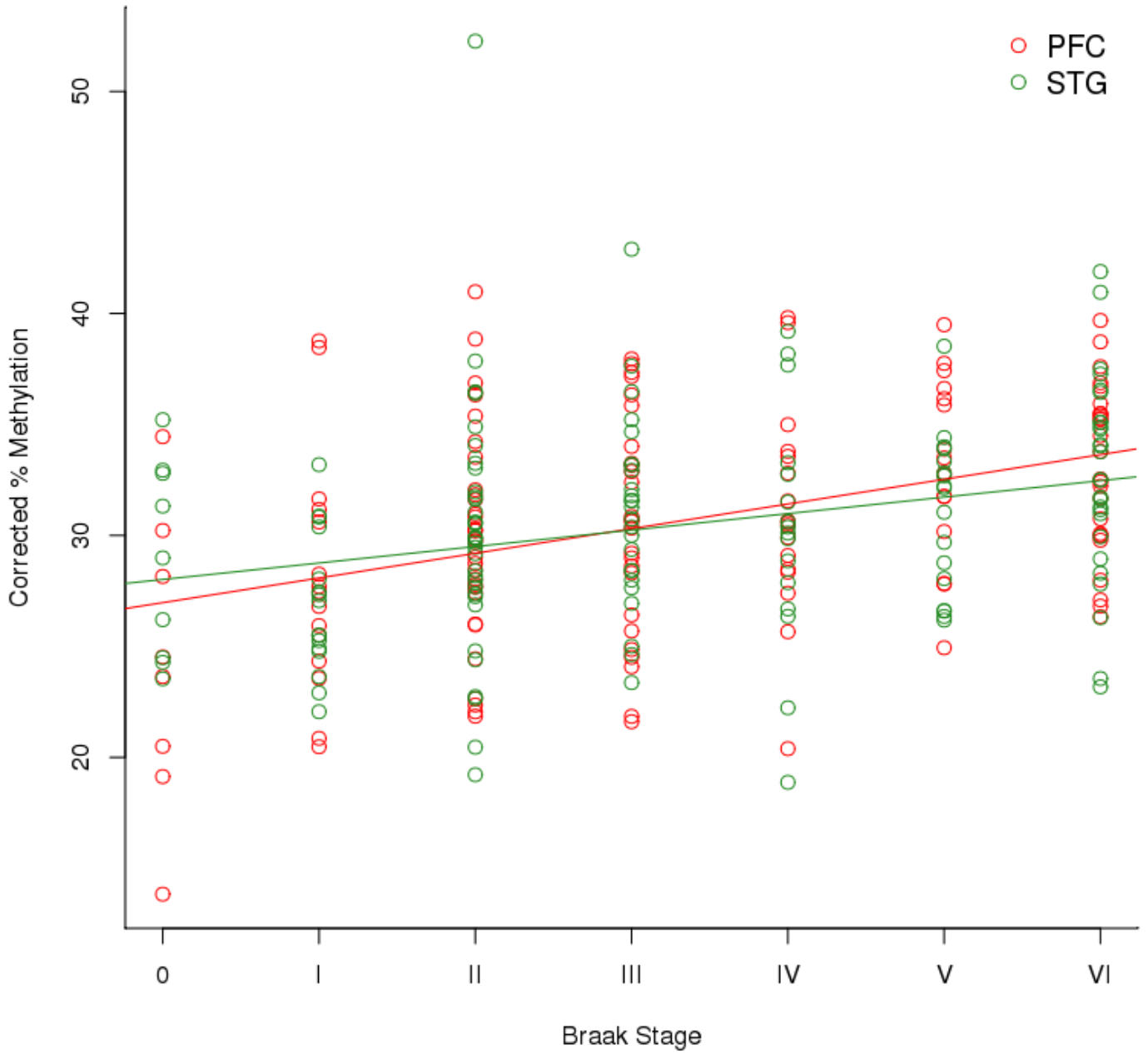

PFC
STG
Corrected % Methylation
20
30
40
50
0
I
II
III
IV
V
VI
Braak Stage

FIGURE 1

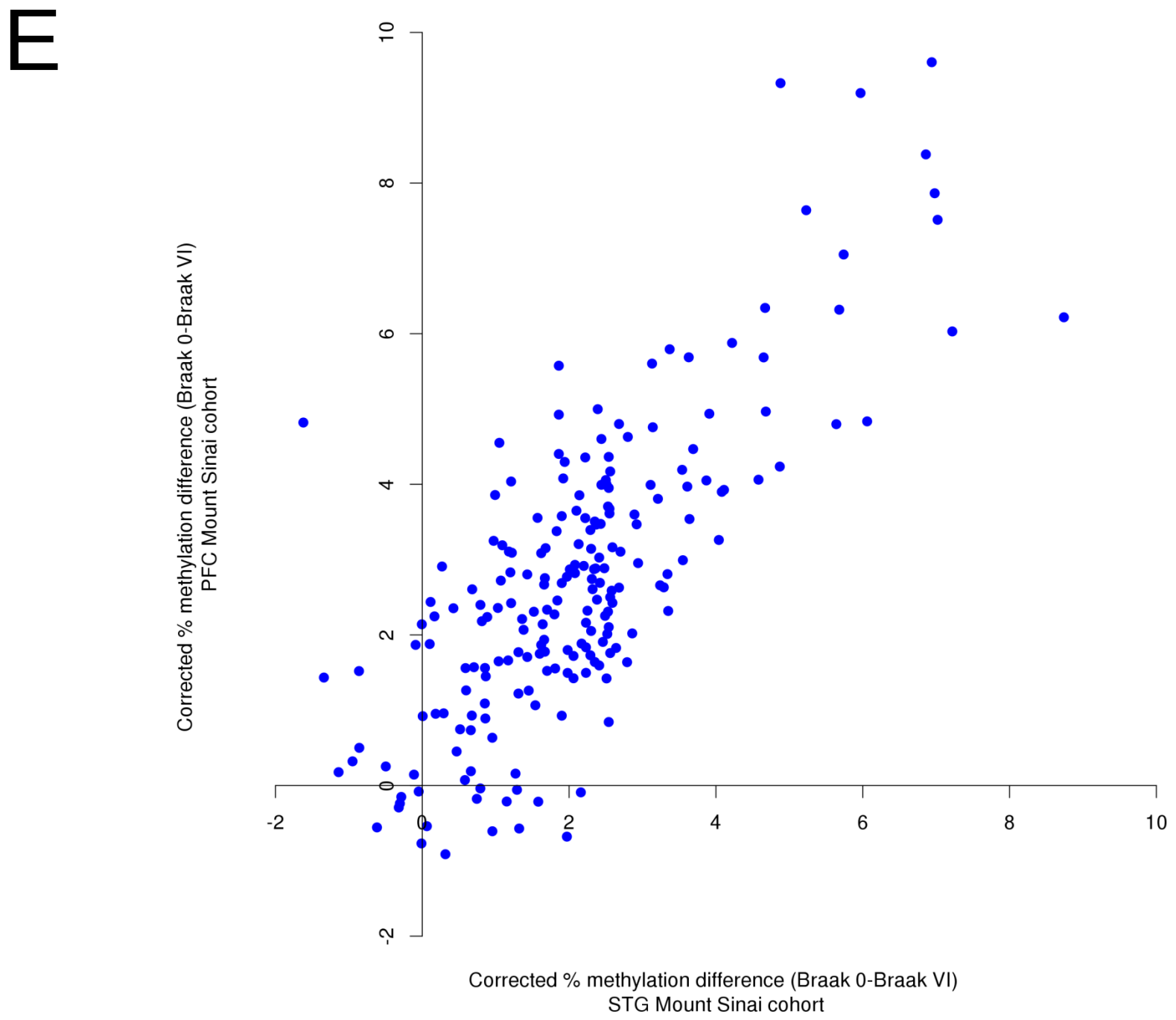

E
Corrected % methylation difference (Braak 0-Braak VI)
PFC Mount Sinai cohort
Corrected % methylation difference (Braak 0-Braak VI)
STG Mount Sinai cohort
-2
0
2
4
6
8
10

FIGURE 2

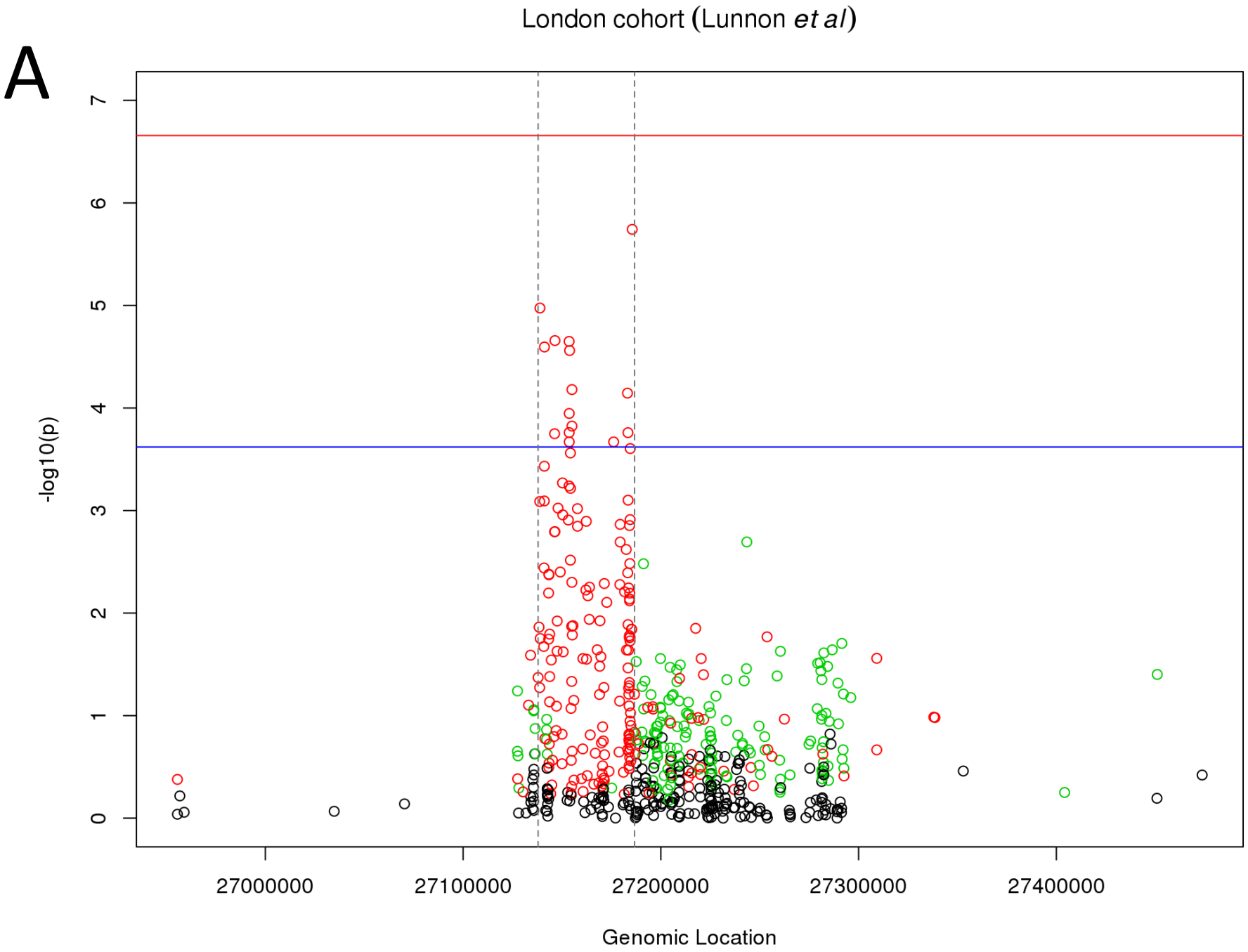

London cohort (Lunnon *et al*)
A
-log10(p)
0
1
2
3
4
5
6
7
27000000
27100000
27200000
27300000
27400000
Genomic Location

FIGURE 2

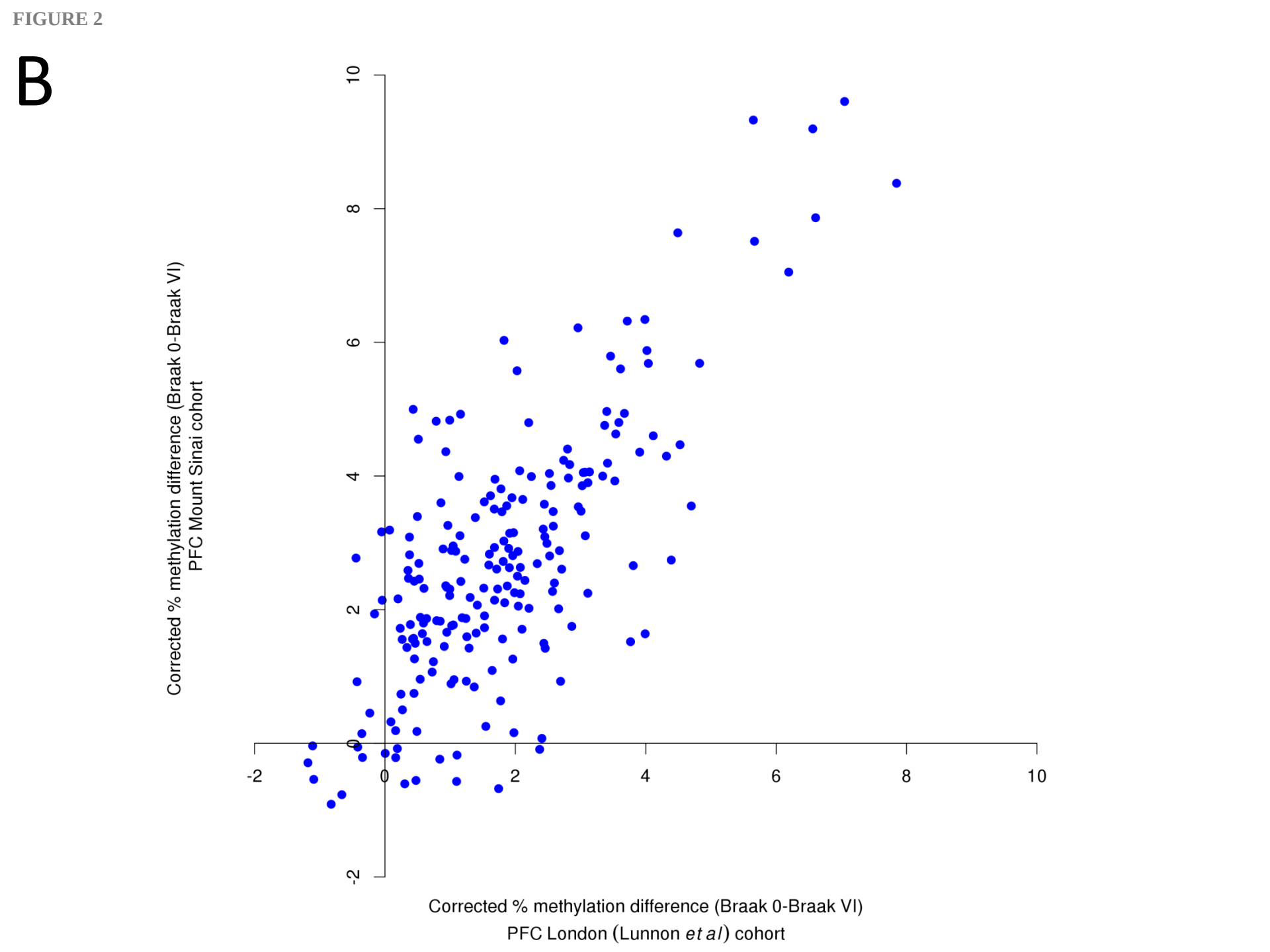

B
Corrected % methylation difference (Braak 0-Braak VI)
PFC Mount Sinai cohort
Corrected % methylation difference (Braak 0-Braak VI)
PFC London (Lunnon et al) cohort
-2
0
2
4
6
8
10

FIGURE 2

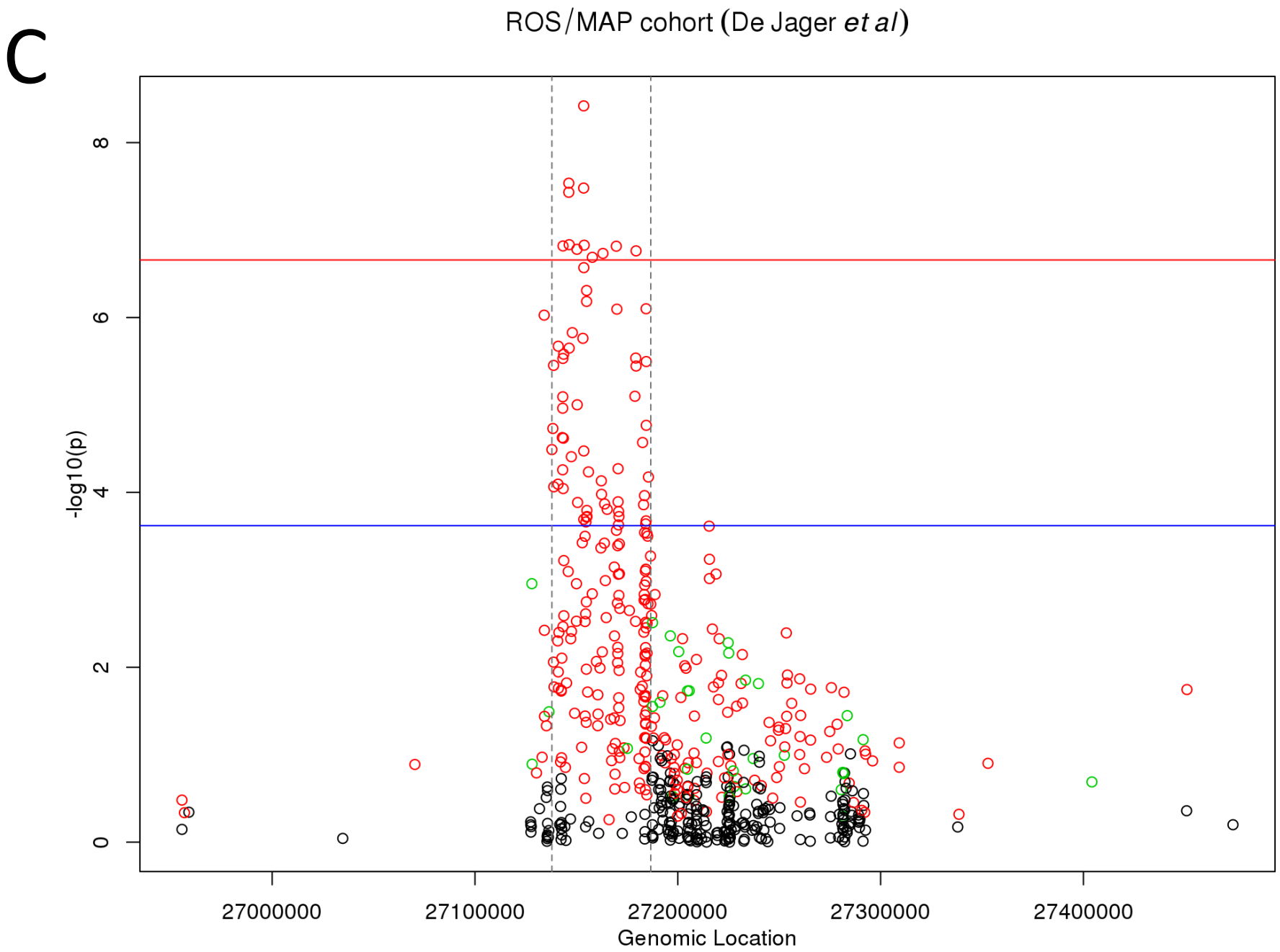

C
ROS/MAP cohort (De Jager et al)
-log10(p)
0
2
4
6
8
27000000
27100000
27200000
27300000
27400000
Genomic Location

FIGURE 2

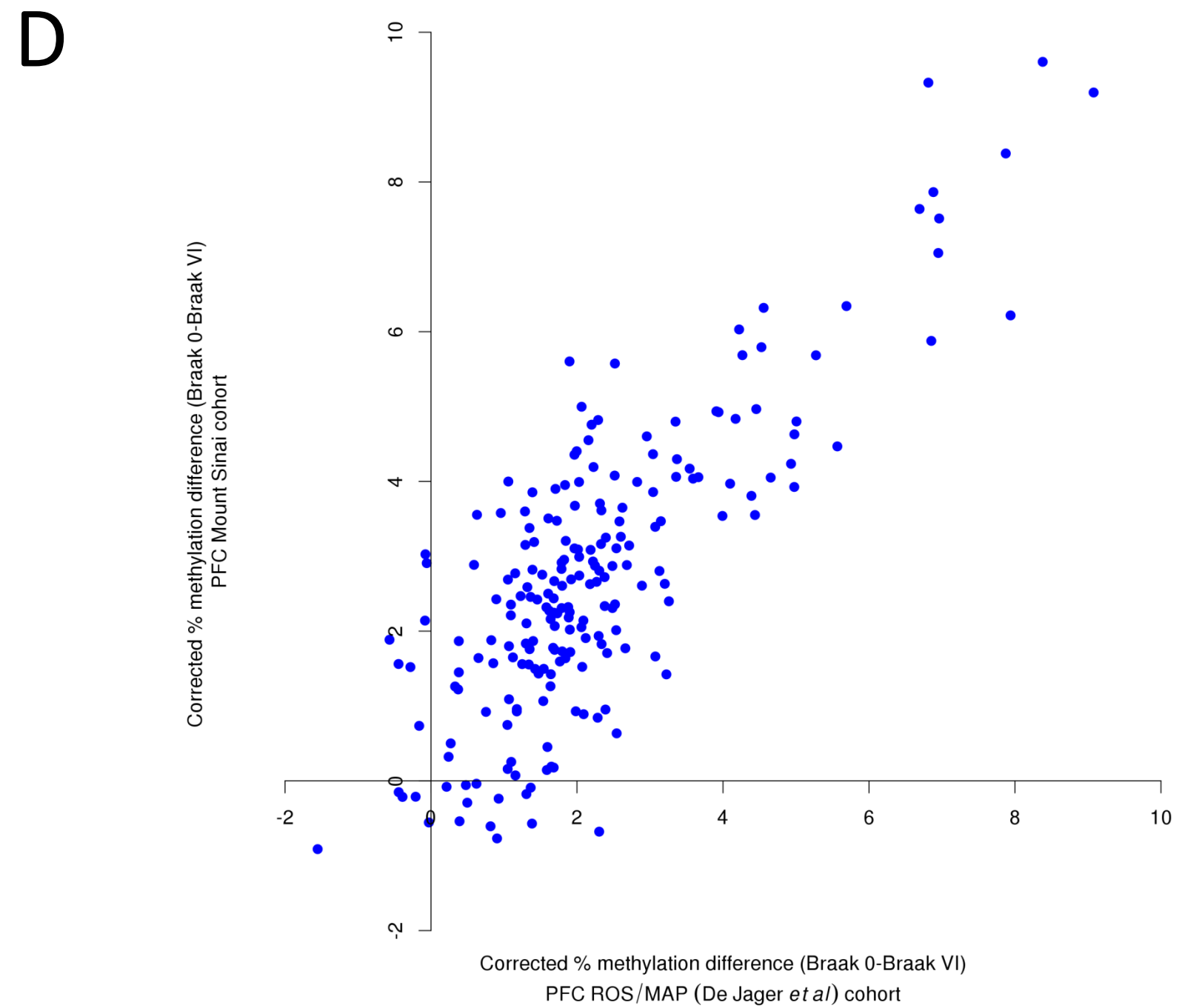

D
Corrected % methylation difference (Braak 0-Braak VI)
PFC Mount Sinai cohort
Corrected % methylation difference (Braak 0-Braak VI)
PFC ROS/MAP (De Jager et al) cohort
-2
0
2
4
6
8
10

FIGURE 2

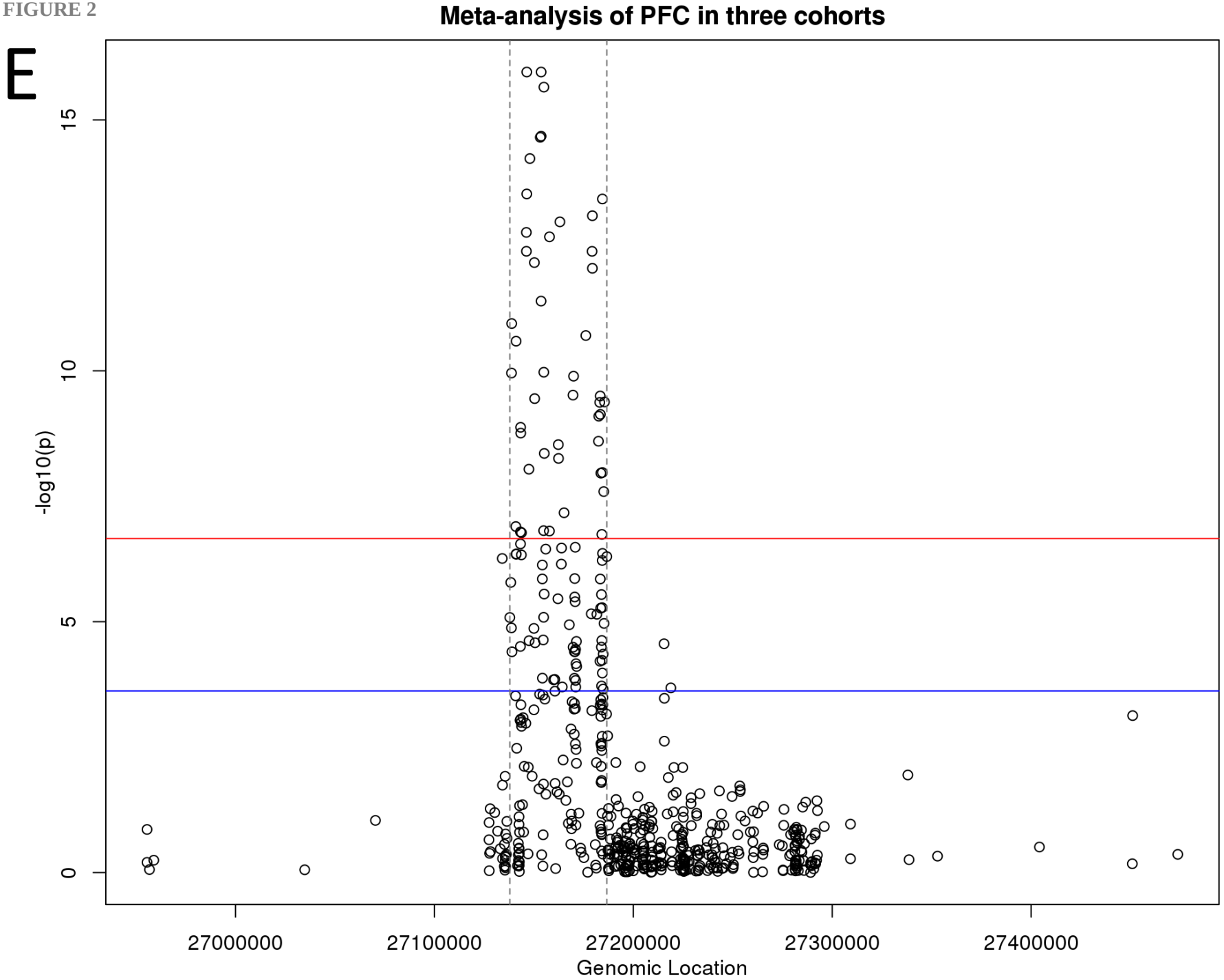

Meta-analysis of PFC in three cohorts
E
-log10(p)
0
5
10
15
27000000
27100000
27200000
27300000
27400000
Genomic Location

FIGURE 2

F

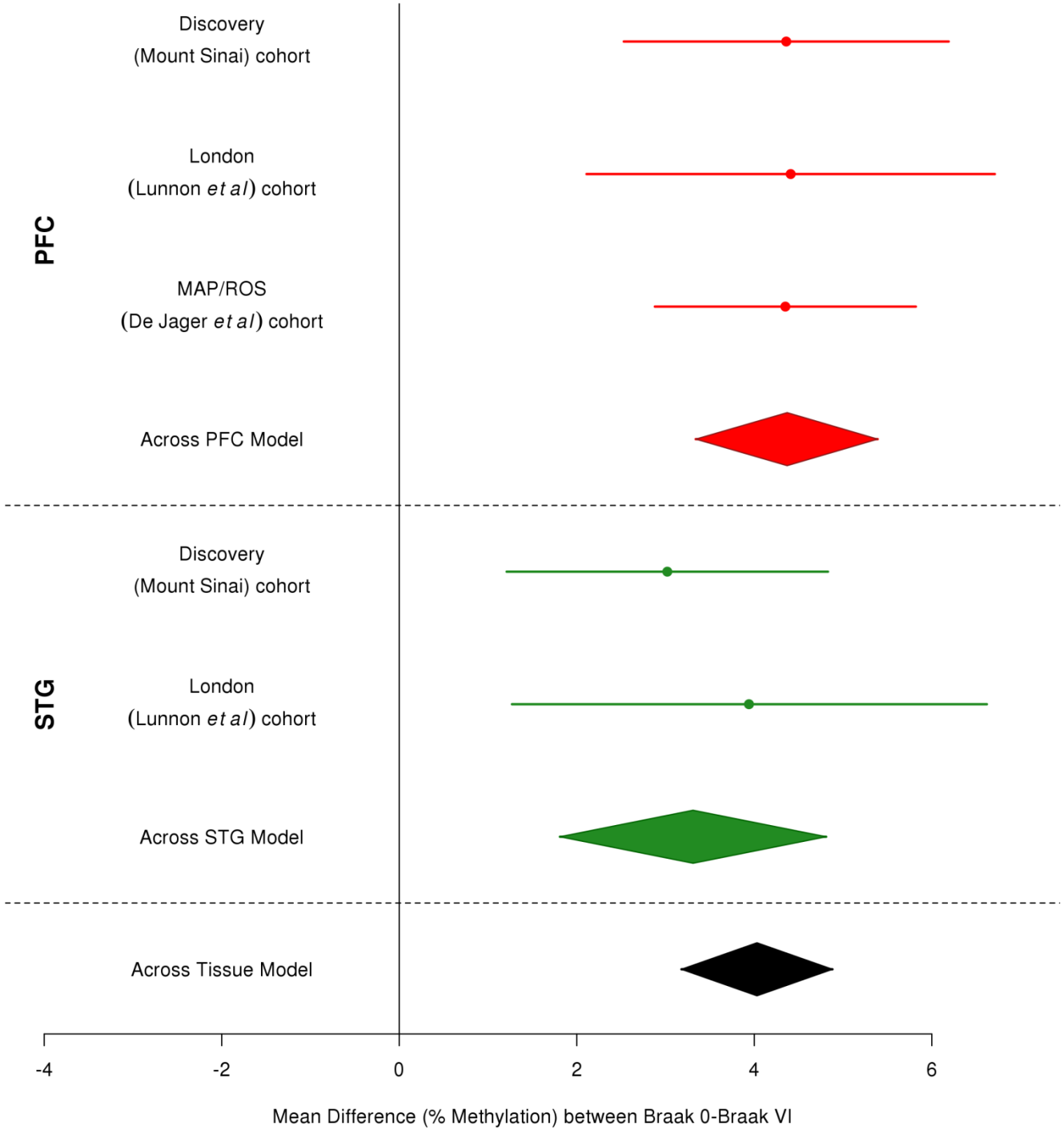

Discovery
(Mount Sinai) cohort
London
(Lunnon et al) cohort
MAP/ROS
(De Jager et al) cohort
PFC
Across PFC Model
Discovery
(Mount Sinai) cohort
London
(Lunnon et al) cohort
STG
Across STG Model
Across Tissue Model
-4
-2
0
2
4
6
Mean Difference (% Methylation) between Braak 0-Braak VI

Table 1: Differentially methylated positions (DMPs) and differentially methylated regions (DMRs) associated with Braak stage in the prefrontal cortex (PFC). (**A**) The ten DMPs in PFC in the discovery (Mount Sinai) cohort that reached experiment-wide significance ($P < 2.2 \times 10^{-7}$) are shown, with annotation to chromosomal location (hg19), up/downstream genes (from GREAT annotation), *P* value from our quantitative association model and corrected DNA methylation difference (Δ) from Braak score 0 – VI (as a %). Also shown is the corresponding information in the matched superior temporal gyrus (STG) samples in the same cohort, and the matched brain regions (PFC, STG) in the London (Lunnon *et al*) cohort, demonstrating a nominally significant difference. A list of the 78 top-ranked PFC DMPs at a more relaxed threshold of $P < 1 \times 10^{-5}$ are given in **Supplementary Table 2.** (**B**) DMRs significantly associated with Braak stage in the PFC. Shown are all significantly associated regions (Sidak-corrected *P* value < 0.05) which contain three or more probes, with chromosomal location (hg19), up/downstream genes, number of probes in the significant region and Sidak corrected *P* value.

A

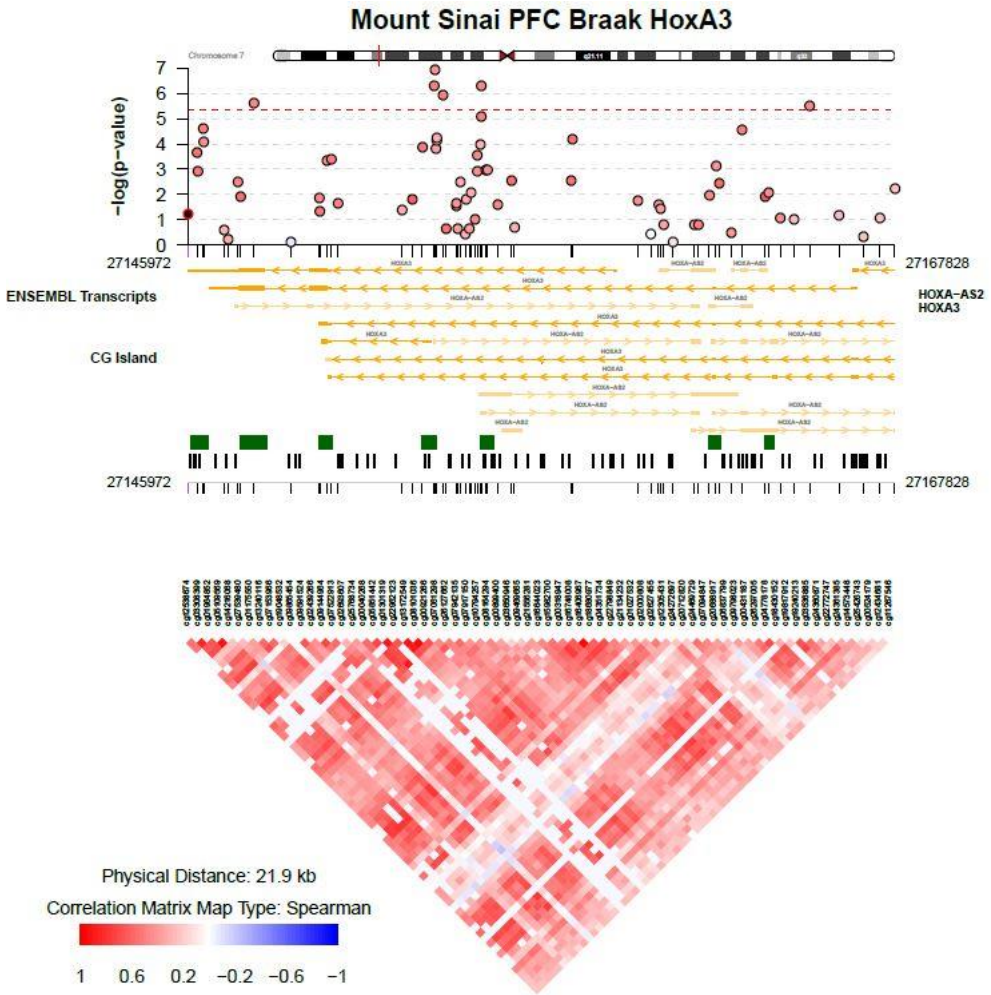
Probe	Location	Illumina Annotation	GREAT Annotation		Discovery (Mount Sinai) cohort				London (Lunnon *et al*) cohort			
					Association with Braak Stage				Association with Braak Stage			
					PFC		STG		PFC		STG	
			Downstream	Upstream	Δ	*P* Value	Δ	*P* Value	Δ	*P* Value	Δ	*P* Value
cg22867816	4:16081205	PROM1	FGFBP2 (-116347)	PROM1 (+4118)	-3.90	9.80E-09	-2.04	5.21E-03	-	-	-	-
cg06977285	7:18127468		HDAC9 (-408457)	PRPS1L1 (-59983)	3.66	2.02E-08	2.68	1.84E-04	-	-	1.88	7.65E-03
cg05783384	2:218843735		RUFY4 (-90242)	TNS1 (-34885)	7.42	4.46E-08	5.55	8.01E-05	3.26	7.76E-03	3.83	6.48E-04
cg07349815	3:123751269		CCDC14 (-70706)	KALRN (-62258)	5.15	6.70E-08	-	-	2.15	0.02	1.83	7.35E-03
cg21806242	11:72532891	ATG16L2	ATG16L2 (+7539)	FCHSD2 (+320414)	8.51	7.02E-08	5.55	4.08E-04	5.22	3.86E-04	4.62	1.10E-03
cg03834767	7:90794392	CDK14	FZD1 (-99390)	CDK14 (+455681)	-4.50	8.13E-08	-	-	-	-	-	-
cg13935577	12:107974897	BTBD11	PWP1 (-104611)	BTBD11 (+262708)	9.11	8.45E-08	5.27	1.49E-03	4.02	5.10E-03	3.73	0.02
cg27078890	11:128457459	ETS1	ETS1 (-23)		4.85	9.86E-08	-	-	2.09	0.02	-	-
cg22962123	7:27153605	HOXA3	HOXA2 (-11176)	HOXA3 (+5608)	7.88	1.20E-07	5.12	2.78E-04	5.62	2.24E-05	5.18	5.21E-04
cg26199857	12:54764265	ZNF385A	GPR84 (-5995)	ZNF385A (+20816)	5.43	1.87E-07	4.44	1.02E-03	2.62	0.03	-	-

B

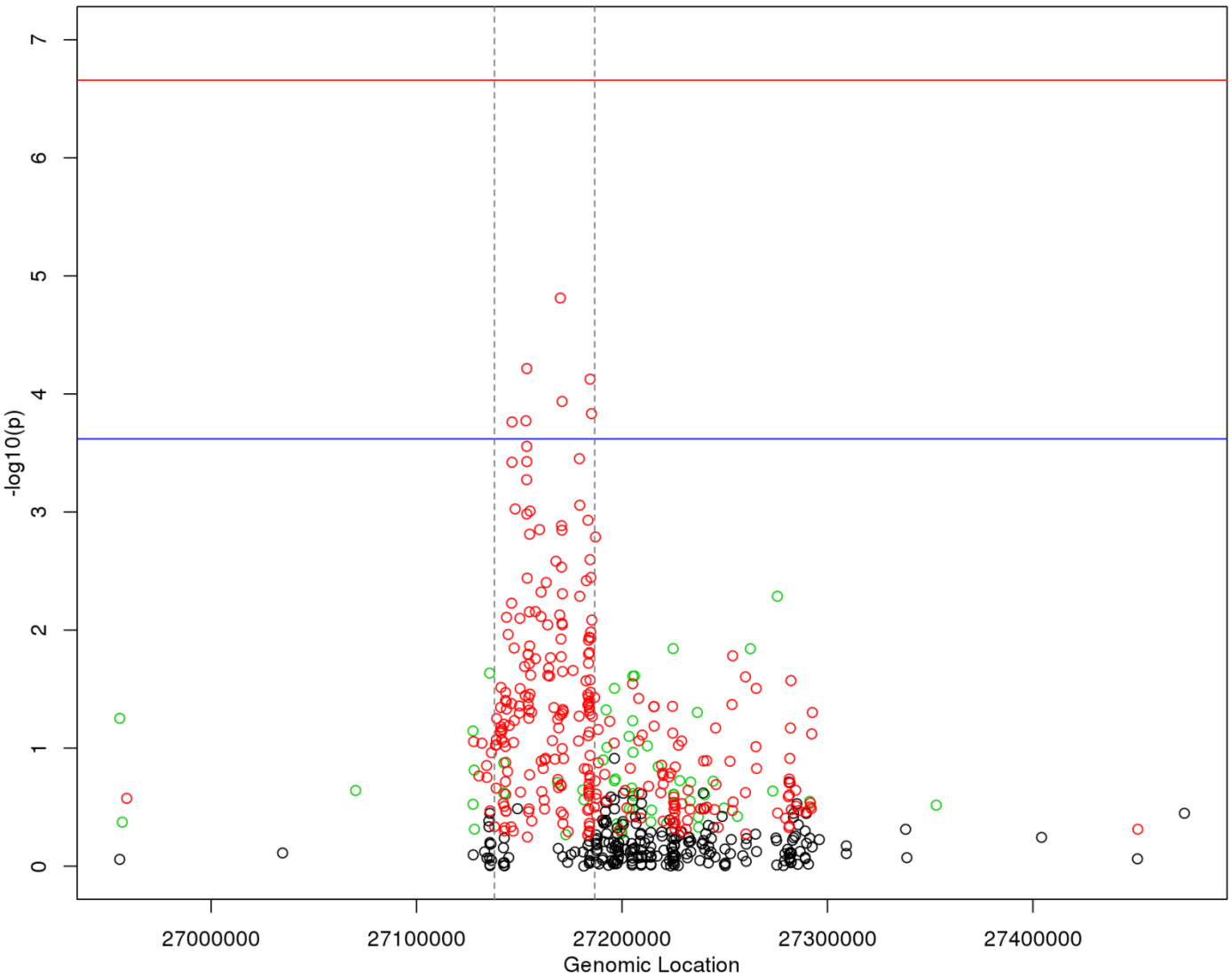
Chr	Start	End	Gene Annotation	GREAT Annotation		Number of Probes	Šidák corrected slk *P* Value
chr11	2321770	2323247	C11ORF21	TSPAN32 (-734)	C11orf21 (+634)	27	3.20E-11
chr7	27153580	27153944	HOXA3	HOXA2 (-11332)	HOXA3 (+5452)	7	1.19E-09
chr7	27154262	27155234	HOXA3	HOXA2 (-12318)	HOXA3 (+4466)	16	4.31E-09
chr7	27169957	27171401	HOXA4	HOXA4 (-261)		21	2.13E-08
chr11	315908	316456	IFITM1 Closest	IFITM1 (+2329)	IFITM3 (+4868)	5	4.02E-08
chr12	58119915	58120237	AGAP2	AGAP2 (+11953)	OS9 (+32172)	6	1.22E-07
chr7	27183133	27184853	HOXA5/HOXA-AS3	HOXA5 (-706)		42	2.19E-06
chr5	78985425	78985900	CMYA5	CMYA5 (-37)		10	2.31E-06
chr19	10736006	10736448	SLC44A2	SLC44A2 (+293)		8	3.68E-06
chr19	39086733	39087186	MAP4K1	MAP4K1 (+21604)	RYR1 (+162490)	4	4.94E-06
chr6	10556147	10556523	GCNT2	GCNT6 (-77658)	GCNT2 (+27746)	3	2.93E-05
chr3	194014592	194015171	GRM2 Closest	CPN2 (+57175)	HES1 (+160948)	4	3.24E-05
chr4	184908351	184909018	STOX2	STOX2 (+82176)	ENPP6 (+230429)	8	3.60E-05
chr7	27145972	27146445	HOXA3	HOXA2 (-3779)		5	4.11E-05
chr17	46388390	46388465	SKAP1	SKAP1 (+119124)	SNX11 (+203508)	3	4.77E-05
chr17	74475240	74475402	RHBDF2	RHBDF2 (+22168)	AANAT (+25888)	5	8.13E-05
chr3	51740741	51741280	GRM2	GRM2 (-75)		6	1.93E-04
chr17	41363502	41364121	NBR1/TMEM106A	TMEM106A (-82)		11	3.04E-04
chr17	43318610	43319371	FMNL1	FMNL1 (+19835)	SPATA32 (+20488)	6	4.51E-04
chr7	158281410	158281613	PTPRN2	PTPRN2 (+98859)		3	4.66E-04
chr13	43565901	43566496	EPSTI1	DNAJC15 (-31140)	TNFSF11 (+417910)	9	4.72E-04
chr20	57582787	57583520	CTSZ Closest	CTSZ (-852)		18	6.82E-04
chr19	3179545	3180035	S1PR4	NCLN (-5808)	S1PR4 (+1054)	4	7.59E-04
chr22	37608611	37608819	SSTR3 Closest	SSTR3 (-353)		3	8.84E-04
chr13	113698408	113699016	MCF2L	F7 (-61409)	MCF2L (+75177)	13	9.15E-04
chr9	34457129	34457500	FAM219A	DNAI1 (-1518)		4	1.05E-03
chr17	75315081	75315567	SEPT9	TNRC6C (-685813)	SEPT9 (+37832)	8	1.28E-03
chr16	29674618	29675214	SPN	SPN (+336)		6	1.77E-03
chr1	55246867	55247408	TTC22	PARS2 (-16951)	DHCR24 (+105753)	5	2.45E-03
chr12	58132558	58133008		AGAP2 (-754)		3	3.00E-03
chr7	27138712	27138974	HOTAIRM1	HOXA1 (-3250)		4	3.19E-03
chr16	67686832	67687392	RLTRP	ACD (+7534)	RLTPR (+8290)	4	3.59E-03
chr12	58129855	58130410	AGAP2	AGAP2 (+1896)	OS9 (+42229)	4	4.42E-03
chr17	19314299	19314618	RNF112	RNF112 (-48)		6	9.80E-03
chr15	40583227	40583422	PLCB2	PLCB2 (+16798)	PAK6 (+51704)	3	0.01922
chr15	38988533	38988860	C15ORF53	THBS1 (-884597)	RASGRP1 (-131690)	4	0.01974
chr16	1482952	1483192	CCDC154 Closest	C16orf91 (-3727)		3	0.02843

Supplementary Figure 1: DNA methylation of probes within the *HOXA3* gene are highly correlated.

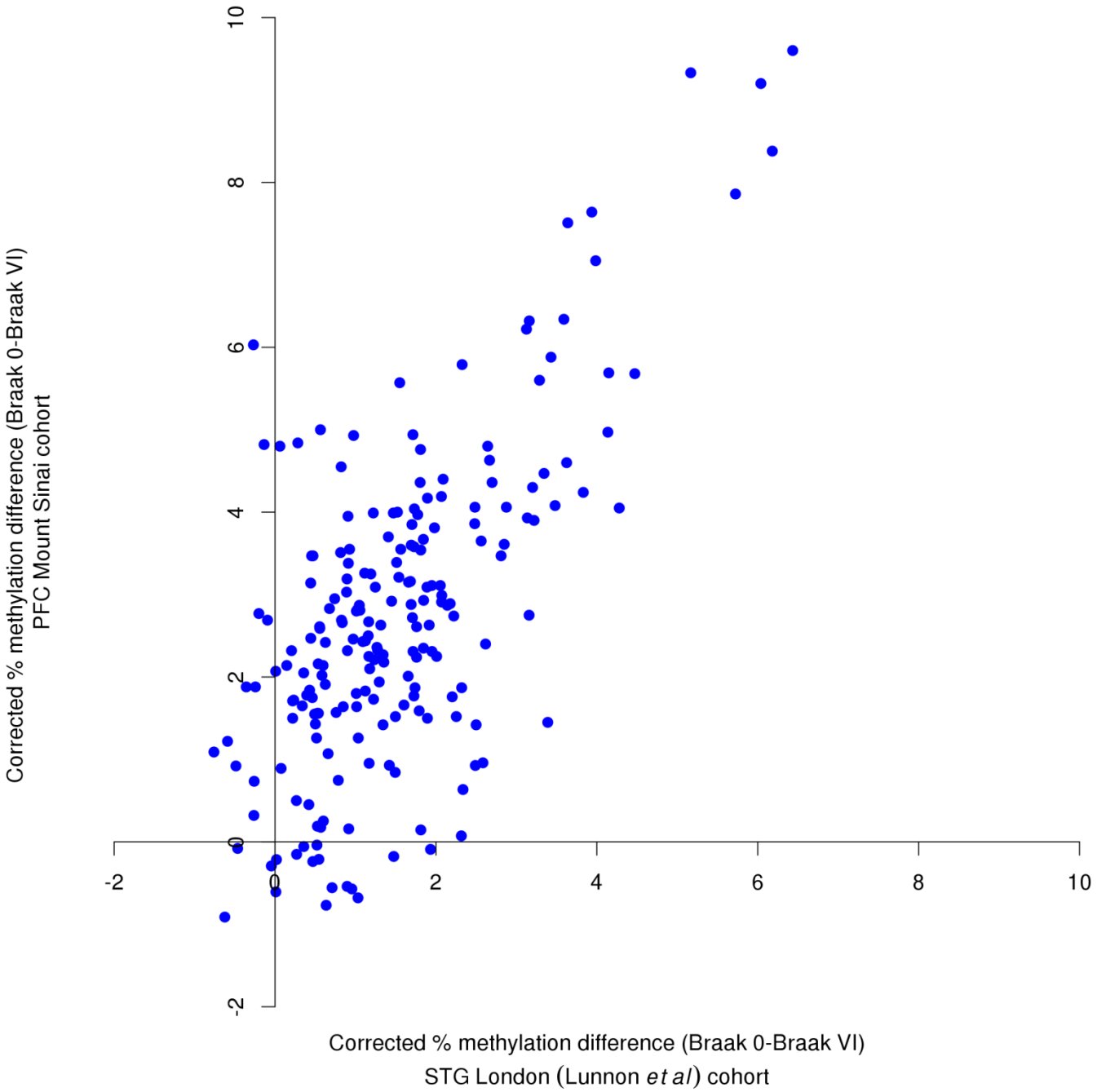
Using the R package *coMET* [17] we saw that the majority of probes within the *HOXA3* gene are highly correlated with each other.



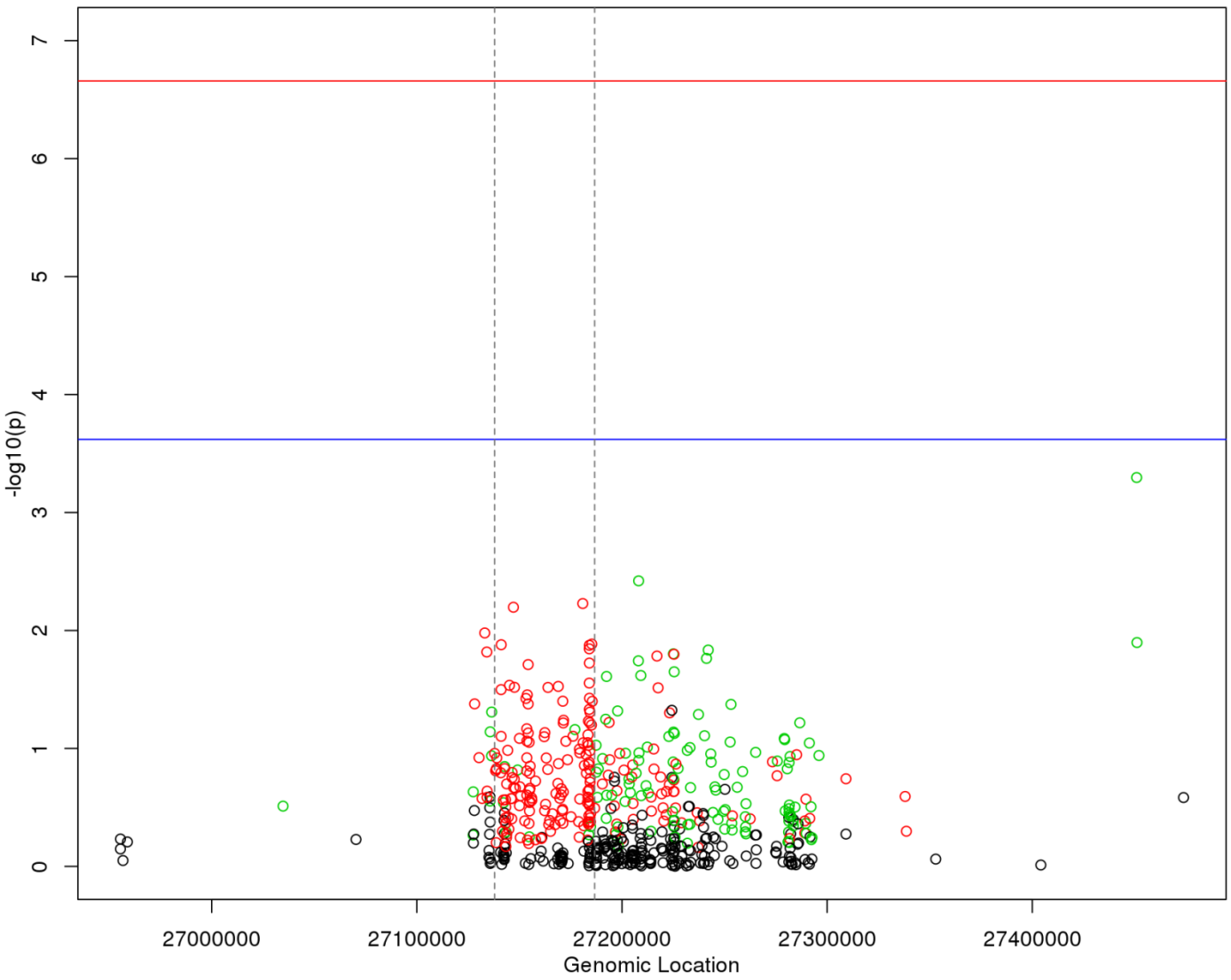
Supplementary Figure 2: A similar pattern of association between DNA methylation in the *HOXA* gene cluster and braak stage was seen in the STG in the London (Lunnon *et al*) cohort. Red circles indicate increased DNA methylation in disease (≥1% between Braak 0 and Braak VI), green circles indicate decreased DNA methylation in disease (≥1% between Braak 0 and Braak VI) and black circles indicate DNA methylation differences <1% between Braak 0 and Braak VI. The red line indicates experiment-wide significance (P = 2.2 X 10^{-7}), whilst the blue line indicates significance after correcting for 208 tests (P = 2.4 X 10^{-4}).



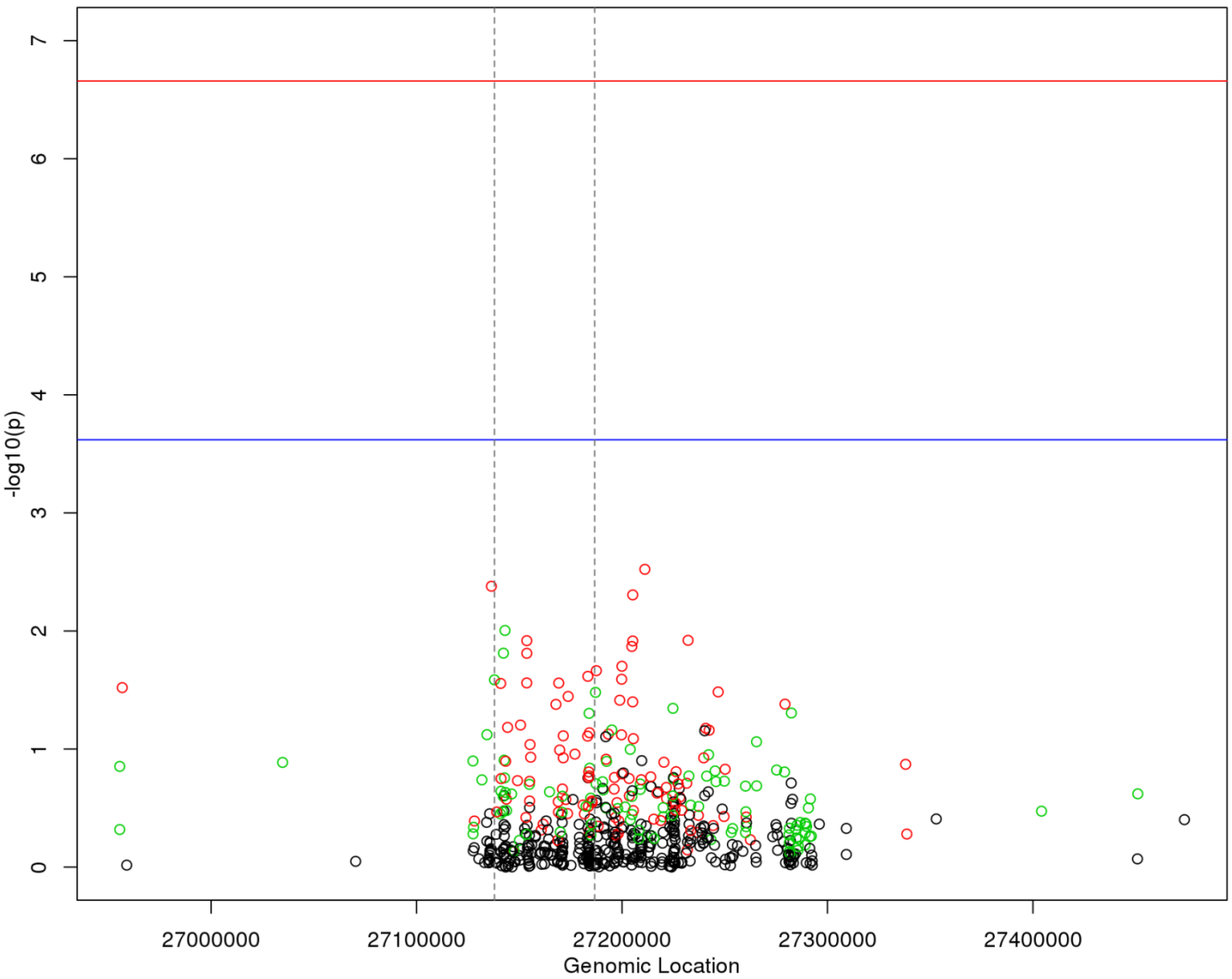
Supplementary Figure 3: Conserved methylation patterns in AD brain are observed across the *HOXA* gene cluster in the STG in the London (Lunnon *et al*) cohort. Methylation patterns in AD STG across the 208 probes in the HOXA gene are highly correlated between the discovery (Mount Sinai) cohort and samples from the London (Lunnon *et al*) cohort (R = 0.68, $P = 1.87 \times 10^{-29}$).



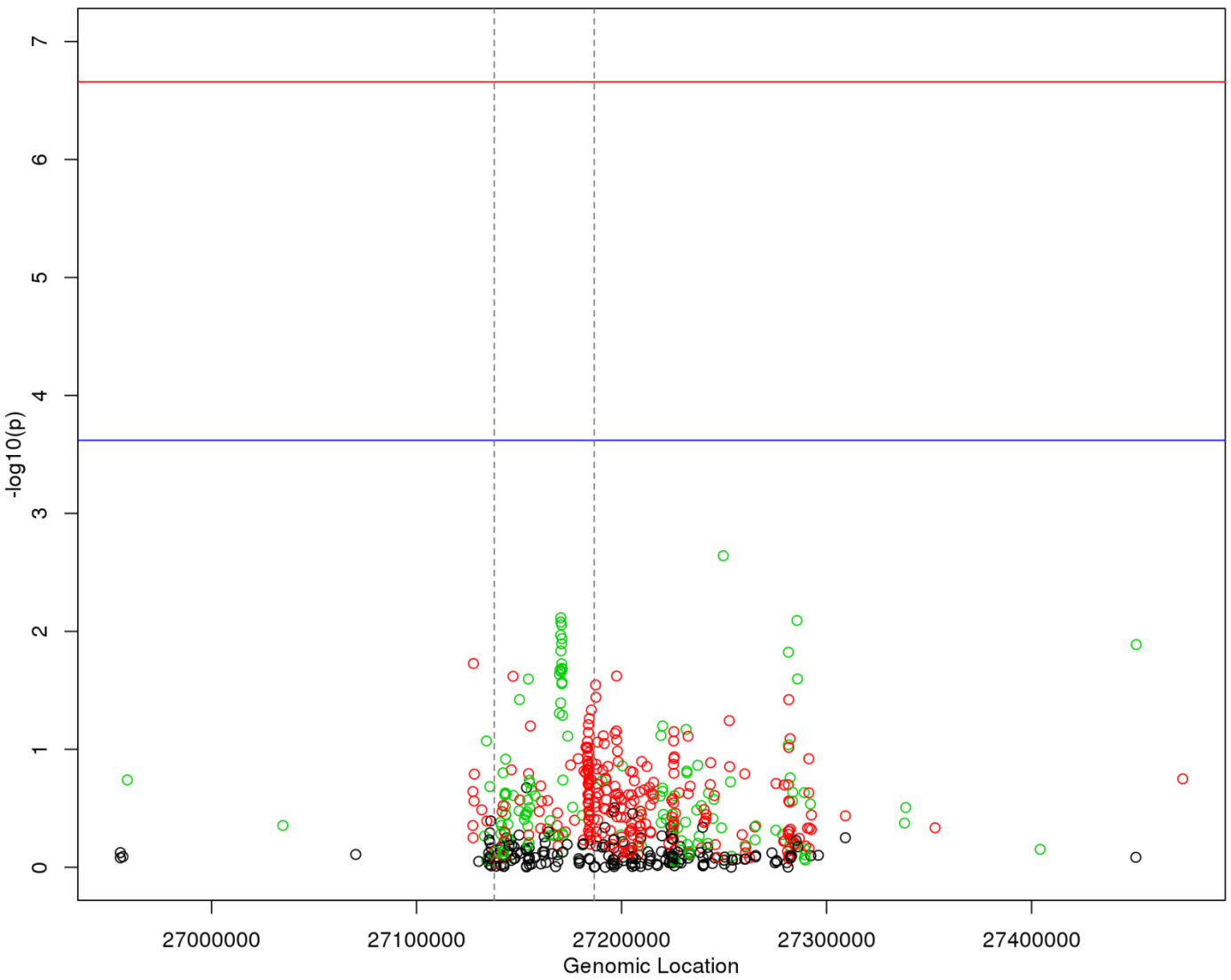
Supplementary Figure 4: No association between DNA hypermethylation in the *HOXA* gene cluster and braak stage was seen in the EC in the London (Lunnon *et al*) cohort. Red circles indicate increased DNA methylation in disease (≥1% between Braak 0 and Braak VI), green circles indicate decreased DNA methylation in disease (≥1% between Braak 0 and Braak VI) and black circles indicate DNA methylation differences <1% between Braak 0 and Braak VI. The red line indicates experiment-wide significance (P = 2.2 X 10^{-7}), whilst the blue line indicates significance after correcting for 208 tests (P = 2.4 X 10^{-4}).



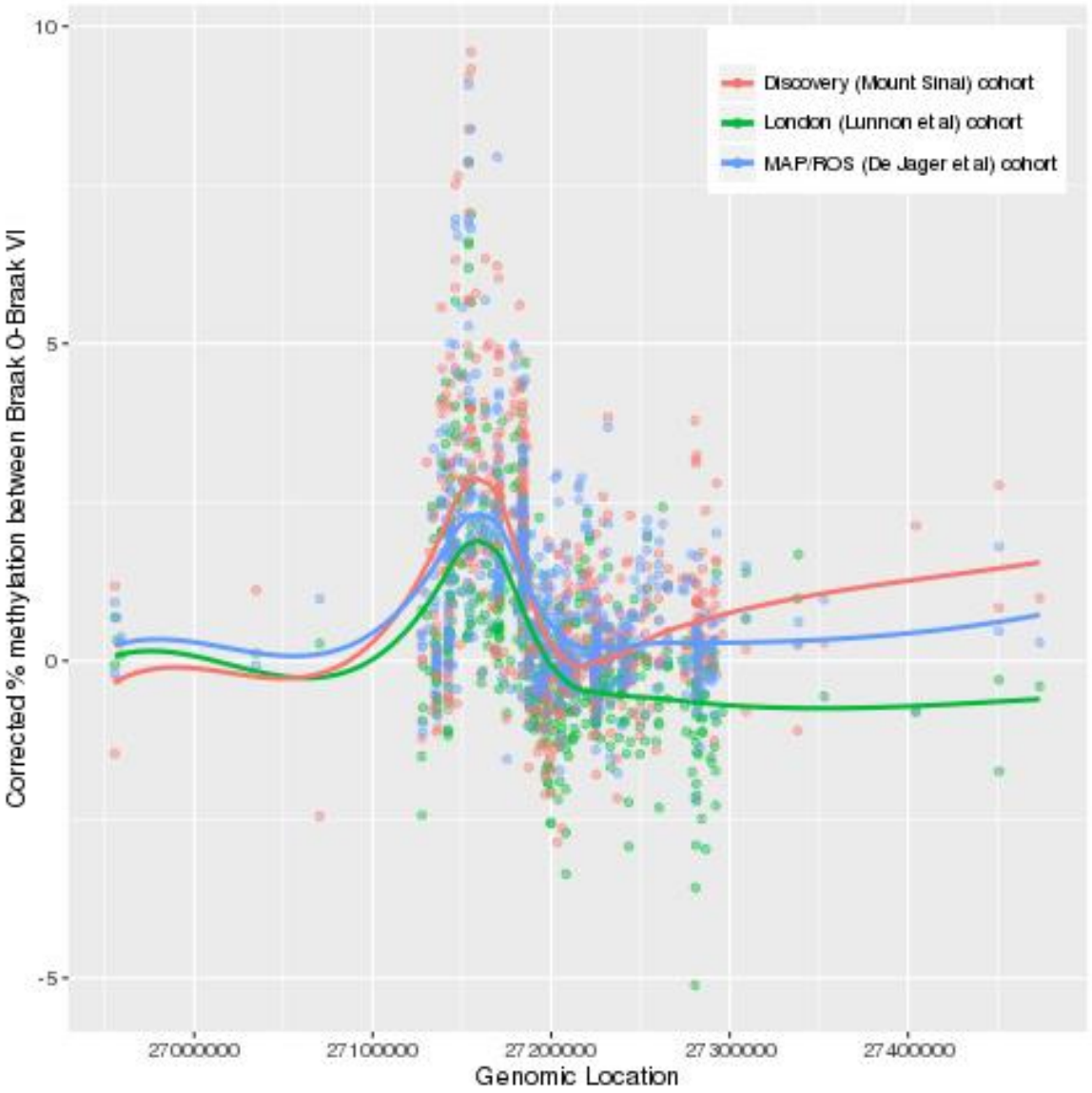
Supplementary Figure 5: No association between DNA hypermethylation in the *HOXA* gene cluster and braak stage was seen in the CER in the London (Lunnon *et al*) cohort. Red circles indicate increased DNA methylation in disease (≥1% between Braak 0 and Braak VI), green circles indicate decreased DNA methylation in disease (≥1% between Braak 0 and Braak VI) and black circles indicate DNA methylation differences <1% between Braak 0 and Braak VI. The red line indicates experiment-wide significance (P = 2.2 X 10^{-7}), whilst the blue line indicates significance after correcting for 208 tests (P = 2.4 X 10^{-4}).



Supplementary Figure 6: No association between DNA hypermethylation in the *HOX* gene cluster and braak stage was seen in pre-mortem blood in the London (Lunnon *et al*) cohort. Red circles indicate increased DNA methylation in disease (≥1% between Braak 0 and Braak VI), green circles indicate decreased DNA methylation in disease (≥1% between Braak 0 and Braak VI) and black circles indicate DNA methylation differences <1% between Braak 0 and Braak VI. The red line indicates experiment-wide significance (P = 2.2 x 10^{-7}), whilst the blue line indicates significance after correcting for 208 tests (P = 2.4 X 10^{-4}).



Supplementary Figure 7: A consistent pattern of effect size was observed across the *HOXA* cluster in all three cohorts



Supplementary Table 1: Demographics for samples used within the study. Abbreviations: SD (standard deviation); PFC (prefrontal cortex); STG (superior temporal gyrus); CER (cerebellum); EC (entorhinal cortex).

	Discovery (Mount Sinai) cohort			**London (Lunnon *et al*) cohort**			**ROS/MAP (De Jager *et al*) cohort**		
	Control	**Middle Stage AD**	**Late Stage**	**Control**	**Middle Stage AD**	**Late Stage**	**Control**	**Middle Stage AD**	**Late Stage**
Number of Cases	60	43	44	29	18	66	151	424	165
Gender (M/F)	32/28	13/30	12/32	13/16	7/11	26/40	75/76	148/276	46/119
Age at Death (±SD)	82.0 (7.56)	82.7 (6.55)	88.0 (7.53)	77.6 (12.80)	88.5 (5.20)	85.4 (8.13)	83.6 (7.19)	88.8 (6.32)	89.8 (5.22)
Braak Score	0-II	III-IV	V-VI	0-II	III-IV	V-VI	0-II	III-IV	V-VI
Brain Regions investigated	PFC, STG			PFC, STG, CER, EC, Blood			PFC		

Supplementary Table 2: Probes associated with Braak stage in the PFC in the discovery (Mount Sinai) cohort at $P < 1 \times 10^{-5}$. Shown are chromosomal location (hg19), up/downstream genes from Illumina and GREAT annotation, *P* Value from our quantitative model, difference (Δ) in corrected DNA methylation from Braak score 0 – VI (as a %). Also shown are the corresponding statistics in the matched STG samples in the same cohort, and the matched brain regions (PFC, STG) in the London (Lunnon *et al*) cohort, demonstrating a nominally significant difference. Abbreviations: PFC (prefrontal cortex); STG (superior temporal gyrus).

Probe	Location	Illumina Annotation	GREAT Annotation		Discovery (Mount Sinai) Cohort				London (Lunnon *et al*) Cohort			
					Association with Braak Stage				Association with Braak Stage			
					PFC		STG		PFC		STG	
					Δ	*P* Value	Δ	*P* Value	Δ	*P* Value	Δ	*P* Value
cg22867816	4:16081205	PROM1	FGFBP2 (-116347)	PROM1 (+4118)	-3.90	9.80E-09	-2.04	5.21E-03	-	-	-	-
cg06977285	7:18127468		HDAC9 (-408457)	PRPS1L1 (-59983)	3.66	2.02E-08	2.68	1.84E-04	-	-	1.88	7.65E-03
cg05783384	2:218843735		RUFY4 (-90242)	TNS1 (-34885)	7.42	4.46E-08	5.55	8.01E-05	3.26	7.76E-03	3.83	6.48E-04
cg07349815	3:123751269		CCDC14 (-70706)	KALRN (-62258)	5.15	6.70E-08	-	-	2.15	0.02	1.83	7.35E-03
cg21806242	11:72532891	ATG16L2	ATG16L2 (+7539)	FCHSD2 (+320414)	8.51	7.02E-08	5.55	4.08E-04	5.22	3.86E-04	4.62	1.10E-03
cg03834767	7:90794392	CDK14	FZD1 (-99390)	CDK14 (+455681)	-4.50	8.13E-08	-	-	-	-	-	-
cg13935577	12:107974897	BTBD11	PWP1 (-104611)	BTBD11 (+262708)	9.11	8.45E-08	5.27	1.49E-03	4.02	5.10E-03	3.73	0.02
cg27078890	11:128457459	ETS1	ETS1 (-23)		4.85	9.86E-08	-	-	2.09	0.02	-	-
cg22962123	7:27153605	HOXA3	HOXA2 (-11176)	HOXA3 (+5608)	7.88	1.20E-07	5.12	2.78E-04	5.62	2.24E-05	5.18	5.21E-04
cg26199857	12:54764265	ZNF385A	GPR84 (-5995)	ZNF385A (+20816)	5.43	1.87E-07	4.44	1.02E-03	2.62	0.03	-	-
cg19802390	1:2840177		MMEL1 (-275749)	ACTRT2 (-97868)	-4.66	3.73E-07	-2.44	6.27E-03	-1.97	5.23E-03	-1.46	0.01
cg02317313	12:122235206	LOC338799	RHOF (-3039)		5.32	4.08E-07	3.76	4.82E-03	3.23	1.83E-03	2.28	0.04
cg06574422	12:111181673	PPP1CC	PPP1CC (-930)		4.21	4.90E-07	2.00	0.05	-	-	-	-
cg01301319	7:27153580	HOXA3	HOXA2 (-11151)	HOXA3 (+5633)	4.87	5.06E-07	3.99	5.33E-04	3.46	5.75E-04	3.83	3.32E-04
cg16406967	7:27155036	HOXA3	HOXA2 (-12607)	HOXA3 (+4177)	7.99	5.12E-07	4.18	0.01	4.84	6.61E-05	4.43	5.23E-04
cg14557202	12:54764371	ZNF385A	GPR84 (-6101)	ZNF385A (+20710)	7.02	5.96E-07	4.40	4.04E-03	3.18	0.02	-	-
cg23377551	12:58130154	AGAP2	AGAP2 (+1874)	OS9 (+42251)	5.68	6.81E-07	4.38	8.28E-04	3.28	6.01E-03	2.87	4.76E-03
cg09642525	16:85320701		GSE1 (-326120)	KIAA0513 (+223884)	2.22	7.76E-07	1.17	0.02	-	-	-	-
cg07315426	7:158281537	PTPRN2	PTPRN2 (+98833)		3.64	8.41E-07	1.59	0.04	-	-	-	-
cg14103343	19:49220223	MAMSTR	MAMSTR (+2754)	FUT2 (+20992)	4.25	8.44E-07	2.23	0.02	1.83	0.02	1.76	0.03
cg16385330	4:151503878	LRBA;MAB21L2	MAB21L2 (+802)		4.84	8.54E-07	3.22	2.95E-03	-	-	-	-
cg08991643	3:194014928		CPN2 (+57128)	HES1 (+160995)	6.83	9.11E-07	4.75	6.38E-04	-	-	-	-
cg07061298	7:27153847	HOXA3	HOXA2 (-11418)	HOXA3 (+5366)	4.87	1.17E-06	3.11	3.63E-03	4.14	2.75E-05	3.56	2.32E-04
cg04682967	22:37608626	SSTR3	SSTR3 (-265)		4.03	1.17E-06	2.45	2.73E-03	1.40	0.03	-	-
cg23120601	15:40583227	PLCB2	PLCB2 (+16895)	PAK6 (+51607)	6.08	1.38E-06	2.72	0.05	3.94	1.86E-03	2.16	0.05
cg02292809	6:17101478		STMND1 (-1010)		-2.35	1.44E-06	-1.27	0.02	-	-	-	-
cg12234455	19:49220235	MAMSTR	MAMSTR (+2742)	FUT2 (+21004)	4.54	1.66E-06	2.80	5.94E-03	1.94	0.02	1.68	0.04
cg02124912	1:202091933	GPR37L1	GPR37L1 (-52)		4.25	1.68E-06	2.94	3.75E-03	2.23	0.04	3.24	2.66E-04
cg17400113	15:93617146	RGMA	RGMA (-55)		-3.05	1.69E-06	-1.93	4.72E-03	-	-	-1.40	0.04
cg23950714	5:176935364	DOK3	PDLIM7 (-10781)	DOK3 (+1493)	-5.76	1.80E-06	-4.31	9.49E-04	-3.48	2.26E-03	-2.60	0.03
cg14769703	10:43818695		RASGEF1A (-114064)	FXYD4 (-48394)	4.39	2.05E-06	2.63	1.34E-03	-	-	-	-

Probe	Location	Illumina Annotation	GREAT Annotation		Discovery (Mount Sinai) Cohort				London (Lunnon *et al*) Cohort			
					Association with Braak Stage				Association with Braak Stage			
					PFC		STG		PFC		STG	
					Δ	*P* Value	Δ	*P* Value	Δ	*P* Value	Δ	*P* Value
cg01463828	8:22446721	PDLIM2	ENSG00000248235 (-65		7.75	2.06E-06	4.63	2.53E-03	4.23	3.20E-03	4.98	1.01E-03
cg17113856	6:32120895	PPT2;PRRT1	PRRT1 (-1167)	PPT2-EGFL8 (-1103)	4.01	2.20E-06	2.75	9.47E-04	2.30	5.52E-03	2.36	1.03E-03
cg26022064	7:98739782	SMURF1	SMURF1 (+1940)	TRRAP (+263670)	4.40	2.26E-06	3.37	1.77E-03	2.59	4.70E-03	2.67	1.93E-03
cg04874795	16:86477638		FOXF1 (-66494)	IRF8 (+545230)	-6.18	2.37E-06	-2.90	0.03	-4.23	3.53E-05	-3.00	3.50E-03
cg05287480	7:27176127		HOXA4 (-5710)	HOXA5 (+7159)	3.30	2.46E-06	1.84	0.02	2.59	2.14E-04	1.46	9.83E-03
cg19048532	7:27148002	HOXA3	HOXA2 (-5573)	HOXA3 (+11211)	6.55	2.52E-06	4.48	9.41E-04	3.85	9.47E-04	3.37	9.47E-03
cg08578641	9:34457440	C9orf25;DNAI1	DNAI1 (-1392)		5.24	2.56E-06	2.81	0.02	2.90	4.06E-03	2.87	5.08E-04
cg09298818	12:111181546	PPP1CC	PPP1CC (-803)		2.11	2.82E-06	1.57	7.15E-04	-	-	-	-
cg01353646	12:58132733	AGAP2	AGAP2 (-705)		5.47	3.20E-06	3.27	9.99E-03	3.23	4.52E-03	-	-
cg14573448	7:27165187	HOXA3	HOXA3 (-5974)	HOXA4 (+5230)	4.28	3.25E-06	2.04	0.02	-	-	-	-
cg12163800	17:74475355	RHBDF2	RHBDF2 (+22133)	AANAT (+25923)	4.42	3.53E-06	2.51	0.02	2.23	2.78E-03	2.71	1.67E-03
cg27228136	22:37608819	SSTR3	SSTR3 (-458)		3.39	3.63E-06	1.58	0.04	-	-	-	-
cg13852561	9:140302117	EXD3	NRARP (-105415)	EXD3 (+15596)	3.92	3.67E-06	2.42	5.64E-03	-	-	-	-
cg08000731	11:316247		IFITM1 (+2395)	IFITM3 (+4802)	8.32	4.12E-06	6.39	5.81E-04	-	-	-	-
cg13714797	9:7103479	KDM4C	KDM4C (+345824)	C9orf123 (+696587)	5.88	4.20E-06	4.00	7.71E-05	-	-	-	-
cg20864214	11:73054121	ARHGEF17	RELT (-33187)	ARHGEF17 (+34788)	5.91	4.39E-06	3.26	0.02	4.92	3.95E-05	4.13	1.02E-03
cg08230957	19:39087186	MAP4K1	MAP4K1 (+21377)	RYR1 (+162717)	4.72	4.51E-06	3.57	2.09E-03	2.87	2.46E-03	2.49	0.01
cg23618477	6:169334872		THBS2 (+319266)	SMOC2 (+493042)	-7.91	4.52E-06	-4.58	5.99E-03	-4.20	9.68E-04	-4.22	1.10E-03
cg17067993	17:38721675	CCR7	CCR7 (+48)		-2.31	4.60E-06	-	-	-	-	-	-
cg21987515	4:153274118	FBXW7	PET112 (-591944)	FBXW7 (+183134)	3.99	4.76E-06	-	-	-	-	-	-
cg08441803	17:46388390	SKAP1	SKAP1 (+119161)	SNX11 (+203471)	3.38	4.82E-06	1.80	0.02	-	-	-	-
cg12155963	7:150147553	GIMAP8	GIMAP8 (-164)		-4.11	4.87E-06	-	-	-	-	-	-
cg14147151	9:138948000	NACC2	NACC2 (-5575)	C9orf69 (+62708)	-4.96	4.94E-06	-3.37	5.56E-03	-	-	-	-
cg13313598	6:136672985	MAP7	BCLAF1 (-61997)	MAP7 (+174624)	4.51	5.15E-06	2.54	0.03	2.16	0.03	2.66	1.64E-03
cg18702012	10:126106555	OAT	OAT (+949)		3.46	5.15E-06	2.37	4.60E-03	-	-	-	-
cg24550149	1:55246954	TTC22	PARS2 (-16768)	DHCR24 (+105936)	4.75	5.26E-06	3.18	3.44E-03	3.48	2.78E-04	-	-
cg14288213	19:18703034	C19orf60	C19orf60 (+3500)	CRLF1 (+14625)	-4.15	5.39E-06	-2.22	9.65E-03	-	-	-	-
cg00795125	7:5409340	TNRC18	TNRC18 (+53836)	SLC29A4 (+86780)	-2.64	5.69E-06	-1.51	8.49E-03	-	-	-	-
cg08924415	8:144902836	PUF60	SCRIB (-5288)	PUF60 (+9192)	3.81	5.77E-06	2.86	3.95E-03	-	-	-	-
cg23347323	19:45912637	ERCC1;CD3EAP	PPP1R13L (-3031)		-2.95	6.00E-06	-1.71	6.28E-03	-	-	-	-
cg09596958	12:58132105	AGAP2	AGAP2 (-77)		7.85	6.45E-06	3.56	0.05	5.86	4.20E-04	4.06	6.91E-03
cg17006136	4:186559412	SORBS2	PDLIM3 (-102751)	SORBS2 (+318393)	-6.02	6.66E-06	-3.08	0.02	-3.03	1.33E-03	-3.50	3.28E-04
cg21730858	6:138806320	NHSL1	ECT2L (-310742)	HEBP2 (+80965)	4.58	7.15E-06	2.42	0.01	-	-	-	-
cg06830450	12:58121004	LOC100130776;AGAP2	AGAP2 (+11024)	OS9 (+33101)	3.89	7.50E-06	2.61	0.01	1.84	0.02	-	-
cg22729726	1:3123854	PRDM16	ARHGEF16 (-247135)	PRDM16 (+138080)	-3.41	7.51E-06	-2.09	0.03	-	-	-1.92	5.41E-03
cg21663431	19:10736355	SLC44A2	SLC44A2 (+422)		6.08	7.92E-06	3.33	0.02	4.26	5.86E-04	3.70	1.45E-03
cg23184252	12:125139714		NCOR2 (-159917)	SCARB1 (+208678)	3.07	7.92E-06	-	-	2.04	8.31E-04	1.05	0.02
cg18680977	7:27155039	HOXA3	HOXA2 (-12610)	HOXA3 (+4174)	8.23	8.15E-06	5.95	1.54E-03	6.04	1.50E-04	5.52	6.22E-04

Probe	Location	Illumina Annotation	GREAT Annotation		Discovery (Mount Sinai) Cohort				London (Lunnon *et al*) Cohort			
					Association with Braak Stage				Association with Braak Stage			
					PFC		STG		PFC		STG	
					Δ	*P* Value	Δ	*P* Value	Δ	*P* Value	Δ	*P* Value
cg15751131	7:140090416	SLC37A3	JHDM1D (-213582)	SLC37A3 (+7932)	-4.05	8.20E-06	-2.16	0.04	-2.35	7.73E-03	-1.57	0.03
cg26681211	11:20385635	HTATIP2	HTATIP2 (+347)		-1.81	8.28E-06	-	-	-	-	-	-
cg19759481	7:27183401	HOXA5	HOXA5 (-115)		3.05	9.05E-06	1.34	0.04	1.60	0.01	-	-
cg14030904	6:149806732	ZC3H12D	ZC3H12D (-536)		-4.34	9.11E-06	-2.72	5.87E-03	-2.07	0.01	-1.84	5.34E-03
cg14795572	12:58131681	AGAP2	AGAP2 (+347)		5.02	9.18E-06	-	-	4.22	7.40E-04	3.21	1.97E-03
cg07737292	16:56892460	MIR138-2	SLC12A3 (-6658)	NUP93 (+128444)	-2.72	9.18E-06	-	-	-1.74	5.18E-03	-	-
cg02411995	16:4851344	ROGDI	SEPT12 (-12823)	ROGDI (+1606)	3.54	9.57E-06	2.81	7.67E-04	-	-	-	-
cg24928023	21:45395775	AGPAT3	TRAPPC10 (-36424)	AGPAT3 (+50887)	5.83	9.79E-06	4.87	2.65E-03	-	-	3.01	8.61E-03
cg24527008	5:177870233	COL23A1	PHYKPL (-210448)	COL23A1 (+147322)	-3.83	9.93E-06	-	-	-1.87	6.11E-03	-	-